

166053

STIC-Biotech/ChemLib

mg

From: Page, Thurman
Sent: Sunday, September 18, 2005 3:33 PM
To: Swope, Sheridan; Chan, Christina; STIC-Biotech/ChemLib
Cc: Woodward, Michael; Richter, Johann; Low, Christopher; Fredman, Jeffrey; Page, Thurman
Subject: RE: 10/719,236

Importance: High

RUSH SEARCH APPROVED

Paul
Schulwitz

RECEIVED
SEP 19 2005
STIC/BIOTECH DIVISION
(STIC)

-----Original Message-----

From: Swope, Sheridan
Sent: Sunday, September 18, 2005 12:56 PM
To: Chan, Christina
Cc: Woodward, Michael; Page, Thurman; Richter, Johann; Low, Christopher; Fredman, Jeffrey
Subject: 10/719,236

May I have this rushed for an allowance?
Pls reply to all, so everyone knows the request has been processed.

For 10/719,236, pls search and interference search:

The translation product of SID 1, residues 232-1164, against the NT and AA data bases.

Note: the translation product's N-terminal and C-terminal residues are:
Met-Asp-Thr-Thr-Val-Pro.....Arg-Arg-Thr-Ser-Lys-Ala

Thank you,

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (uspro)

Query Match	100.0%;	Score 1636;	DB 7;	Length 311;
Best Local Similarity	100.0%;	Pred. No. 6.4e-168;		
Matches 311; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

```
QY      1 MDTTVPFSLAEIQGLHODEFRRCRLRDKGLFYLTDGCLTDTELKSAKDVIDFEEHGE 60
DB      1 MDTTVPFSLAEIQGLHODEFRRCRLRDKGLFYLTDGCLTDTELKSAKDVIDFEEHGE 60
QY      61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSSYSDYSMCYSMGTADNLPSPGDFERIMT 120
DB      61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSSYSDYSMCYSMGTADNLPSPGDFERIMT 120
QY      121 QYFDRQYTSRAVARREVLRATGTEPDGVEAFLDCEPLLRFPRYPQVPEHRSABEQPLRM 180
DB      121 QYFDRQYTSRAVARREVLRATGTEPDGVEAFLDCEPLLRFPRYPQVPEHRSABEQPLRM 180
QY      181 APHYDLSMTVLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVKA 240
DB      181 APHYDLSMTVLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVKA 240
QY      241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB      241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY      301 NYVNIIRRTSKA 311
DB      301 NYVNIIRRTSKA 311
```

RESULT 2

AAU17775

ID AAU17775 standard; protein, 311 AA.

XX AC AAM17775;

DT 17-FEB-1998 (first entry)

DE S. clavuligerus wild type expandase.

XX KM Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
KM thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
KM semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
KM cephalosporin ring; DOAC.

XX OS Streptomyces clavuligerus.

XX PN MO9720053-A2.

XX PD 05-JUN-1997.

XX PF 27-NOV-1996; 96WO-EP005358.

XX PR 27-NOV-1995; 95EP-00203259.

XX PR 27-NOV-1995; 95US-0007564P.

XX PA (KONN) GIST-BROCADES BV.

XX PA (UYOX-) UNITV OXFORD.

XX PI Sutherland JD, Bovenberg RAL, Van Der Laan JM;

XX DR WPI; 1997-310608/28.

XX PT Expandase enzymes with mutation(s) to alter substrate specificity - used
XX particularly for production of 7-amino:desacetoxy:cephalosporanic acid.

XX PS Disclosure; Fig 1; 21pp; English.

XX CC This sequence represents the wild type expandase from Streptomyces
XX clavuligerus. The DNA encoding this sequence was subjected to site
XX directed mutagenesis to create the mutant expandase sequences of the
XX invention (see AAM34166-W34177). The mutant expandase sequences have an
XX altered substrate specificity compared to the wild type expandase.
XX Expandase catalyses the expansion of the 5-membered thiazolidine ring of
XX penicillin N to the 6-membered dihydrothiazine ring of DOAC. The mutant
XX expandase gene is used particularly for the production of 7-
XX aminodesacetoxycephalosporanic acid (7-ADCA). 7-ADCA is one of the key
XX intermediates that is used in the preparation of semi-synthetic

CC cephalosporins. The mutant expandases have improved ability to expand
CC penicillin G into a cephalosporin ring in vitro and/or in vivo where
CC other penicillins such as penicillin N and isopenicillin N can act as
CC competing substrates

SQ Sequence 311 AA;

Query Match 99.7%; Score 1634; DB 2; Length 311;
Best Local Similarity 99.7%; Pred. No. 1.1e-167;
Matches 310; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 MDTTVPFSLAEIQGLHODEFRRCRLRDKGLFYLTDGCLTDTELKSAKDVIDFEEHGE 60
DB      1 MDTTVPFSLAEIQGLHODEFRRCRLRDKGLFYLTDGCLTDTELKSAKDVIDFEEHGE 60
QY      61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSSYSDYSMCYSMGTADNLPSPGDFERIMT 120
DB      61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSSYSDYSMCYSMGTADNLPSPGDFERIMT 120
QY      121 QYFDRQYTSRAVARREVLRATGTEPDGVEAFLDCEPLLRFPRYPQVPEHRSABEQPLRM 180
DB      121 QYFDRQYTSRAVARREVLRATGTEPDGVEAFLDCEPLLRFPRYPQVPEHRSABEQPLRM 180
QY      181 APHYDLSMTVLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVKA 240
DB      181 APHYDLSMTVLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVGGQVKA 240
QY      241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB      241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY      301 NYVNIIRRTSKA 311
DB      301 NYVNIIRRTSKA 311
```

RESULT 3

AAU11044

ID AAU11044 standard; protein, 311 AA.

XX AC AAU11044;

XX DT 12-MAR-2002 (first entry)

XX DE Streptomyces clavuligerus penicillin N expandase protein sequence.

XX KM Penicillin N expandase; ring-expanding; penicillin G;
KM phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid; cephalosporin;
KM cephalaxin; penicillin V.

XX OS Streptomyces clavuligerus.

XX PN WC200185951-A1.

XX PD 15-NOV-2001.

XX PF 09-MAY-2001; 2001MO-GB002047.

XX PR 09-MAY-2000; 2000GB-00011185.

XX PA (ACSD-) ACS DOBPAR UK LTD.

XX PI Johnson RI, Newbert RW;

XX DR WPI; 2002-075247/10.

XX PR N-SDB; AAS17241.

XX PT New enzyme useful for ring-expanding penicillin G to produce phenylacetyl
XX -7-amino desacetoxycephalosporanic acid, comprises penicillin expandase
XX having increased specificity for substrates such as penicillin G.

XX PS Claim 4; Fig 1; 42pp; English.

CC The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not the
CC natural substrate of the unmodified expandase. The invention is useful
CC for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA (amino
CC desacetoxyccephalosporanic acid) and the phenylacetyl side chain is
CC removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful as a
CC starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as a
CC probe. The modification of the penicillin expandase enhances the activity
CC of the enzyme such as penicillin N expandase for penicillin G or V as a
CC substrate. The modified expandase has enhanced catalytic activity or
CC increased specificity for another substrate such as penicillin G. The
CC present protein sequence represents the penicillin N expandase enzyme of
CC the invention

XX SQ Sequence 311 AA;

Query Match 99.9%; Score 1634; DB 5; Length 311;

Best Local Similarity 99.7%; Pred. No. 1.1e-167; Mismatches 0; Indels 0; Gaps 0;

Matches 310; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTVPTFSLAELOOGLHODEFRRLDKGLFYLTDCGLTDTLKSANDVIDPFEHSGE 60

DB 1 MDTVPFSLAELOOGLHODEFRRLDKGLFYLTDCGLTDTLKSANDVIDPFEHSGE 60

QY 61 AEKRAVTSVPVPMRGFTGLESESTAOITNTGSDYSCYMGATADNLFPSPGDFERITWT 120

DB 61 AEKRAVTSVPVPMRGFTGLESESTAOITNTGSDYSCYMGATADNLFPSPGDFERITWT 120

QY 121 QYFDRQYASRAVAEVLRAATGTEPDGVEAFLLCEPLRFYFPQVPEHRSABEQPIRM 180

DB 121 QYFDRQYASRAVAEVLRAATGTEPDGVEAFLLCEPLRFYFPQVPEHRSABEQPIRM 180

QY 181 APHYDLISWVTLTIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGSGGYKA 240

DB 181 APHYDLISWVTLTIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGSGGYKA 240

QY 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFDWIGG 300

DB 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFDWIGG 300

QY 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

DB 301 NYVNIRRTSKA 311

XX 09-MAY-2000; 2000GB-00011185.

PR (ACSD-) ACS DOBFAR UK LTD.

PA Johnson RI, Newbert RW;

XX WPI, 2002-075247/10.

DR New enzyme useful for ring-expanding penicillin G to produce phenylacetyl

XX -7-amino desacetoxyccephalosporanic acid, comprises penicillin expandase

PT having increased specificity for substrates such as penicillin G.

XX Claim 8; Page: 42pp; English.

XX The present invention relates to new penicillin expandases modified to

CC improve the ring-expanding activity for a substrate which is not the

CC natural substrate of the unmodified expandase. The invention is useful

CC for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA (amino

CC desacetoxyccephalosporanic acid) and the phenylacetyl side chain is

CC removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful as a

CC starting point for the production of a range of semi-synthetic

CC cephalosporins, most notably cephalixin. The polynucleotide of the

CC invention is useful for producing a primer e.g. a PCR primer and as a

CC probe. The modification of the penicillin expandase enhances the activity

CC of the enzyme such as penicillin N expandase for penicillin G or V as a

CC substrate. The modified expandase has enhanced catalytic activity or

CC increased specificity for another substrate such as penicillin G. This

CC sequence is one of a collection of Streptomyces clavuligerus mutants,

CC AAU1046-AAU1054 and AAU1521-AAU1537, used in the invention to improve

CC the ring-expanding activity of expandase. Note: The present protein

CC sequence is not shown in the specification but is derived from the

CC Streptomyces clavuligerus wild-type penicillin N expandase sequence given

CC in Figure 1 (see AAU1044)

XX SQ Sequence 311 AA;

Query Match 99.8%; Score 1632; DB 5; Length 311;

Best Local Similarity 99.4%; Pred. No. 1.7e-167; Mismatches 0; Indels 0; Gaps 0;

Matches 309; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTVPTFSLAELOOGLHODEFRRLDKGLFYLTDCGLTDTLKSANDVIDPFEHSGE 60

DB 1 MDTVPFSLAELOOGLHODEFRRLDKGLFYLTDCGLTDTLKSANDVIDPFEHSGE 60

QY 61 AEKRAVTSVPVPMRGFTGLESESTAOITNTGSDYSCYMGATADNLFPSPGDFERITWT 120

DB 61 AEKRAVTSVPVPMRGFTGLESESTAOITNTGSDYSCYMGATADNLFPSPGDFERITWT 120

QY 121 QYFDRQYASRAVAEVLRAATGTEPDGVEAFLLCEPLRFYFPQVPEHRSABEQPIRM 180

DB 121 QYFDRQYASRAVAEVLRAATGTEPDGVEAFLLCEPLRFYFPQVPEHRSABEQPIRM 180

QY 181 APHYDLISWVTLTIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGSGGYKA 240

DB 181 APHYDLISWVTLTIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGSGGYKA 240

QY 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFDWIGG 300

DB 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFDWIGG 300

```

XX Streptomyces clavuligerus penicillin N expandase I158V mutant.
DE
XX Penicillin N expandase; ring-expanding; penicillin G;
XX phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid; cephalosporin;
XX cephalalexin; penicillin V; mutant; mutein.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 158
FT /note= "Wild-type Leu substituted by Val"
XX
PN W0200185951-A1.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-GB002047.
XX
XX 09-MAY-2000; 2000GB-00011185.
XX
XX (ACSD-) ACS DOBPAR UK LTD.
XX Johnson RI, Newbert RW;
XX
XX WPI; 2002-075247/10.
XX
XX New enzyme useful for ring-expanding penicillin G to produce phenylacetyl
XX -7-amino desacetoxycephalosporanic acid, comprises penicillin expandase
XX having increased specificity for substrates such as penicillin G.
XX
XX Claim 6; Page; 42pp; English.
XX
XX The present invention relates to new penicillin expandases modified to
XX improve the ring-expanding activity for a substrate which is not the
XX natural substrate of the unmodified expandase. The invention is useful
XX for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA (amino
XX desacetoxycephalosporanic acid) and the phenylacetyl side chain is
XX removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful as a
XX starting point for the production of a range of semi-synthetic
XX cephalosporins, most notably cephalalexin. The polynucleotide of the
XX invention is useful for producing a primer e.g. a PCR primer and as a
XX probe. The modification of the penicillin expandase enhances the activity
XX of the enzyme such as penicillin N expandase for penicillin G or V as a
XX substrate. The modified expandase has enhanced catalytic activity or
XX increased specificity for another substrate such as penicillin G. This
XX sequence is one of a collection of Streptomyces clavuligerus mutants,
XX AAU11046-AAU11054 and AAU11521-AAU11537, used in the invention to improve
XX the ring-expanding activity of expandase. Note: The present protein
XX sequence is not shown in the specification but is derived from the
XX Streptomyces clavuligerus wild-type penicillin N expandase sequence given
XX in figure 1 (see AAU11044)
XX
SQ Sequence 311 AA;
Query March 99.7%; Score 1631; DB 5; Length 311;
Best Local Similarity 99.4%; Pred. No. 2.2e-167;
Matches 309; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTVPTFSLAELOOGLHODEFFRRCLNDKGLFYLTDCGLDTELEKSAKDIVIDFFEHGSE 60
Db 1 MDTVPTFSLAELOOGLHODEFFRRCLNDKGLFYLTDCGLDTELEKSAKDIVIDFFEHGSE 60
QY 61 AEKRAVTSVPVTFMRGFTGLESESTAOITNTGYSYSDYSCYSGMTADNLFPSGDFERIWT 120
Db 61 AEKRAVTSVPVTFMRGFTGLESESTAOITNTGYSYSDYSCYSGMTADNLFPSGDFERIWT 120
QY 121 QYPRORATASRAVAREVLRATGTBPDGVEAFLLCEPLTARRYPPOVPEHRSABEQPLRM 180
Db 121 QYPRORATASRAVAREVLRATGTBPDGVEAFLLCEPLTARRYPPOVPEHRSABEQPLRM 180
QY 181 APHYDLSTMVTLIOOTPCANGFVSLQAEVGAFPTDLPYRPDAVLVFCGAIATLVTGQVKA 240

```

```

Db 181 APHYDLSTMVTLIOOTPCANGFVSLQAEVGAFPTDLPYRPDAVLVFCGAIATLVTGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFELRPNADFTFTSVLAECGFDSLDDETFATFOWMIG 300
Db 241 PRHHVAAPRRDQIAGSSRTSSVFFELRPNADFTFTSVLAECGFDSLDDETFATFOWMIG 300
QY 301 NYNIRRTSKA 311
Db 301 NYNIRRTSKA 311
RESULT 6
ID AAU11051 standard; protein; 311 AA.
XX
XX AAU11051;
XX
XX 12-MAR-2002 (first entry)
XX
DE Streptomyces clavuligerus penicillin N expandase I305M mutant.
XX
XX Penicillin N expandase; ring-expanding; penicillin G;
XX phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid; cephalosporin;
XX cephalalexin; penicillin V; mutant; mutein.
XX
OS Streptomyces clavuligerus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 305
FT /note= "Wild-type Ile substituted by Met"
XX
XX W0200185951-A1.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-GB002047.
XX
XX 09-MAY-2000; 2000GB-00011185.
XX
XX (ACSD-) ACS DOBPAR UK LTD.
XX Johnson RI, Newbert RW;
XX
XX WPI; 2002-075247/10.
XX
XX New enzyme useful for ring-expanding penicillin G to produce phenylacetyl
XX -7-amino desacetoxycephalosporanic acid, comprises penicillin expandase
XX having increased specificity for substrates such as penicillin G.
XX
XX Claim 8; Page; 42pp; English.
XX
XX The present invention relates to new penicillin expandases modified to
XX improve the ring-expanding activity for a substrate which is not the
XX natural substrate of the unmodified expandase. The invention is useful
XX for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA (amino
XX desacetoxycephalosporanic acid) and the phenylacetyl side chain is
XX removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful as a
XX starting point for the production of a range of semi-synthetic
XX cephalosporins, most notably cephalalexin. The polynucleotide of the
XX invention is useful for producing a primer e.g. a PCR primer and as a
XX probe. The modification of the penicillin expandase enhances the activity
XX of the enzyme such as penicillin N expandase for penicillin G or V as a
XX substrate. The modified expandase has enhanced catalytic activity or
XX increased specificity for another substrate such as penicillin G. This
XX sequence is one of a collection of Streptomyces clavuligerus mutants,
XX AAU11046-AAU11054 and AAU11521-AAU11537, used in the invention to improve
XX the ring-expanding activity of expandase. Note: The present protein
XX sequence is not shown in the specification but is derived from the
XX Streptomyces clavuligerus wild-type penicillin N expandase sequence given
XX in figure 1 (see AAU11044)

```

SQ Sequence 311 AA;

Query Match 99.7%; Score 1631; DB 5; Length 311;
Best Local Similarity 99.4%; Pred. No. 2.2e-167;
Matches 309; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTVPFSLAELOQGLHODEFRCLRDKGLFYLTDCGLTDTLKSADIVIDPFHSGE 60
DB 1 MDTVPFSLAELOQGLHODEFRCLRDKGLFYLTDCGLTDTLKSADIVIDPFHSGE 60
QY 61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSDYSVMCYMGTAADNLFPSSGDFERIMT 120
DB 61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSDYSVMCYMGTAADNLFPSSGDFERIMT 120
QY 121 QYFDROYTASRAVAAREVIRATGTEPDGVEAFLDCEPLIRFRYPQVPEHRSABEQPIRM 180
DB 121 QYFDROYTASRAVAAREVIRATGTEPDGVEAFLDCEPLIRFRYPQVPEHRSABEQPIRM 180
QY 181 APHYDLMSWTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLMSWTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVINIRRTSKA 311
DB 301 NYVINIRRTSKA 311

RESULT 7
AAU11053
ID AU11053 standard; protein; 310 AA.
XX AAU11053;
XX
DT 12-MAR-2002 (first entry)
XX
DE S. clavuligerus penicillin N expandase 1 residue deletion at C-terminus.
XX
KW Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid; cephalosporin;
KW cephalalexin; penicillin V; mutant; mutein.
XX
OS Streptomycetes clavuligerus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 310 /note= "Protein is truncated at this residue"
XX
XX WO200185951-A1.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001MO-GB002047.
XX
XX 09-MAY-2000; 2000GB-00011185.
XX
XX (ACSD-) ACS DOBFAR UK LTD.
XX
XX Johnson RI, Newbert RW;
XX
XX WPI; 2002-075247/10.
XX
XX
XX New enzyme useful for ring-expanding penicillin G to produce phenylacetyl
XX -7-amino desacetoxyccephalosporanic acid, comprising penicillin expandase
XX having increased specificity for substrates such as penicillin G.
XX
XX Disclosure; Page; 42pp; English.
XX
XX The present invention relates to new penicillin expandases modified to

CC improve the ring-expanding activity for a substrate which is not the
CC natural substrate of the unmodified expandase. The invention is useful
CC for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA (amino
CC desacetoxyccephalosporanic acid) and the phenylacetyl side chain is
CC removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful as a
CC starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalalexin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as a
CC probe. The modification of the penicillin expandase enhances the activity
CC of the enzyme such as penicillin N expandase for penicillin G or V as a
CC substrate. The modified expandase has enhanced catalytic activity or
CC increased specificity for another substrate such as penicillin G. This
CC sequence is one of a collection of Streptomycetes clavuligerus mutants,
CC AU11046-AU11054 and AU11521-AU11537, used in the invention to improve
CC the ring-expanding activity of expandase. Note: The present protein
CC sequence is not shown in the specification but is derived from the
CC Streptomycetes clavuligerus wild-type penicillin N expandase sequence given
CC in figure 1 (see AU11044)

SQ Sequence 310 AA;

Query Match 99.6%; Score 1630; DB 5; Length 310;
Best Local Similarity 99.7%; Pred. No. 2.8e-167;
Matches 309; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTVPFSLAELOQGLHODEFRCLRDKGLFYLTDCGLTDTLKSADIVIDPFHSGE 60
DB 1 MDTVPFSLAELOQGLHODEFRCLRDKGLFYLTDCGLTDTLKSADIVIDPFHSGE 60
QY 61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSDYSVMCYMGTAADNLFPSSGDFERIMT 120
DB 61 AEKRAVTSVPVPTMRGFTGLSESTAOITNTGSDYSVMCYMGTAADNLFPSSGDFERIMT 120
QY 121 QYFDROYTASRAVAAREVIRATGTEPDGVEAFLDCEPLIRFRYPQVPEHRSABEQPIRM 180
DB 121 QYFDROYTASRAVAAREVIRATGTEPDGVEAFLDCEPLIRFRYPQVPEHRSABEQPIRM 180
QY 181 APHYDLMSWTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLMSWTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVINIRRTSK 310
DB 301 NYVINIRRTSK 310

RESULT 8
AAU11525
ID AU11525 standard; protein; 310 AA.
XX AAU11525;
XX
DT 12-MAR-2002 (first entry)
XX
DE S. clavuligerus penicillin N expandase K310 deletion mutant.
XX
KW Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino desacetoxyccephalosporanic acid; cephalosporin;
KW cephalalexin; penicillin V; mutant; mutein.
XX
OS Streptomycetes clavuligerus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 310 /note= "Protein is truncated at this residue"
XX
XX WO200185951-A1.
XX
XX

PD 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-GB002047.
XX
XX 09-MAY-2000; 2000GB-0001185.
XX
XX (ACSD-) ACS DOBFAR UK LTD.
XX
XX Johnson RI, Newbert RW;
XX WPI; 2002-075247/10.
XX
XX New enzyme useful for ring-expanding penicillin G to produce phenylacetyl
PT -7-amino deacetoxycephalosporanic acid, comprises penicillin expandase
PT having increased specificity for substrates such as penicillin G.
XX
XX Example; Page; 42pp; English.
XX
XX The present invention relates to new penicillin expandases modified to
CC improve the ring-expanding activity for a substrate which is not the
CC natural substrate of the unmodified expandase. The invention is useful
CC for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA (amino
CC deacetoxycephalosporanic acid) and the phenylacetyl side chain is
CC removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful as a
CC starting point for the production of a range of semi-synthetic
CC cephalosporins, most notably cephalixin. The polynucleotide of the
CC invention is useful for producing a primer e.g. a PCR primer and as a
CC probe. The modification of the penicillin expandase enhances the activity
CC of the enzyme such as penicillin N expandase for penicillin G or V as a
CC substrate. The modified expandase has enhanced catalytic activity or
CC increased specificity for another substrate such as penicillin G. This
CC sequence is one of a collection of Streptomyces clavuligerus mutants,
CC AAU11046-AAU11054 and AAU11521-AAU11537, used in the invention to improve
CC the ring-expanding activity of expandase. Note: The present protein
CC sequence is not shown in the specification but is derived from the
CC Streptomyces clavuligerus wild-type penicillin N expandase sequence given
CC in figure 1 (see AAU11044)
XX
XX Sequence 310 AA;

Query Match 99.6%; Score 1630; DB 5; Length 310;
Best Local Similarity 99.7%; Pred. No. 2.8e-167;
Matches 309; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTTVPFISLAELQOGLHODEFRRCIRDKGLFYLTDCGLTDTLTKSAKDIVIDFFEHGSE 60
DB 1 MDTTVPFISLAELQOGLHODEFRRCIRDKGLFYLTDCGLTDTLTKSAKDIVIDFFEHGSE 60
QY 61 AEKRAVTSVPVPTMRGFTGLESESTAOITNTGSSYSDYSCMSGTAADNLFPSSGDFERIMT 120
DB 61 AEKRAVTSVPVPTMRGFTGLESESTAOITNTGSSYSDYSCMSGTAADNLFPSSGDFERIMT 120
QY 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
DB 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
QY 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
DB 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
QY 181 APHYDLSMVTLLIOQTCCANGFVSLQAEVGAFTDLPYRPDAVLPFGAIALTLVGGQVKA 240
DB 181 APHYDLSMVTLLIOQTCCANGFVSLQAEVGAFTDLPYRPDAVLPFGAIALTLVGGQVKA 240
QY 241 PRHHVAAPRRDOJAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDOJAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
QY 301 NYVNIRRTSK 310
DB 301 NYVNIRRTSK 310

RESULT 9
AAP93215
ID AAP93215 standard; protein; 311 AA.
XX

AC AAP93215;
XX
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-MAR-1990 (first entry)
XX
XX Deacetoxycephalosporin C synthetase.
DE
XX Deacetoxycephalosporin C synthetase; DAOCs; cephalosporin; penicillin N;
KM antibiotics.
KM
XX Streptomyces clavuligerus.
OS
XX EP341892-A.
PN
XX 15-NOV-1989.
PD
XX
XX 09-MAY-1988; 88US-00192273.
PF
XX 09-MAY-1988; 88US-00192273.
PR
XX (ELIL) LILLY & CO ELI.
PA
XX Ingolia TD, Kovacevic S, Miller JR, Skatrud PL;
PI WPI; 1989-334231/46.
DR N-PSDB; AAN92262.
XX
XX New recombinant DNA encoding de-acetoxy-cephalosporin C synthetase - for
PT increasing or inducing cephalosporin synthesis in microorganisms or in
PT vitro.
PT
XX
XX Claim 2; Page 26-27; 46pp; English.
PS
XX
XX The sequence encodes deacetoxycephalosporin C synthetase (DAOCs) which
CC catalyses expansion of penicillin N. Cephalosporium strains transformed
CC with vectors carrying the gene will produce antibiotics more efficiently.
CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to
CC correct OS field.)
CC
XX
XX Sequence 311 AA;

Query Match 99.6%; Score 1630; DB 1; Length 311;
Best Local Similarity 99.7%; Pred. No. 2.9e-167;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDTTVPFISLAELQOGLHODEFRRCIRDKGLFYLTDCGLTDTLTKSAKDIVIDFFEHGSE 60
DB 1 MDTTVPFISLAELQOGLHODEFRRCIRDKGLFYLTDCGLTDTLTKSAKDIVIDFFEHGSE 60
QY 61 AEKRAVTSVPVPTMRGFTGLESESTAOITNTGSSYSDYSCMSGTAADNLFPSSGDFERIMT 120
DB 61 AEKRAVTSVPVPTMRGFTGLESESTAOITNTGSSYSDYSCMSGTAADNLFPSSGDFERIMT 120
QY 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
DB 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
QY 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
DB 121 QYPRQYTASRAVAREVLATGTEPDGVEAFIDCEPILRRFYFPQVPEHRSABEQPLRM 180
QY 181 APHYDLSMVTLLIOQTCCANGFVSLQAEVGAFTDLPYRPDAVLPFGAIALTLVGGQVKA 240
DB 181 APHYDLSMVTLLIOQTCCANGFVSLQAEVGAFTDLPYRPDAVLPFGAIALTLVGGQVKA 240
QY 241 PRHHVAAPRRDOJAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDOJAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
QY 301 NYVNIRRTSKA 311
DB 301 NYVNIRRTSKA 311

RESULT 10
AAM34168

```

ID  AAW34168 standard; protein; 311 AA.
XX
AC  AAW34168;
XX
DT  17-FEB-1998 (first entry)
XX
DE  Expandase mutant R74Q.
XX
KM  Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
KM  thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
KM  semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
KM  cephalosporin ring; DOAC; mutant.
XX
OS  Streptomyces clavuligerus.
OS  Synthetic.
XX
FH  Key
FT  Misc-difference 74
FT  Location/Qualifiers
XX  /note= "Arg to Gln mutation"
XX
PN  MO9720053-A2.
XX
PD  05-JUN-1997.
XX
PF  27-NOV-1996; 96WO-EP005358.
XX
PR  27-NOV-1995; 95EP-00203259.
PR  27-NOV-1995; 95US-0007564P.
XX
PA  (KONN ) GIST-BROCADES BV.
XX  (UYOX-) UNIV OXFORD.
XX
PI  Sutherland JD, Bovenberg RAL, Van Der Laan JM;
XX  WPI; 1997-310608/28.
XX
DR  WPI; 1997-310608/28.
XX
PT  Expandase enzymes with mutation(s) to alter substrate specificity - used
PT  particularly for production of 7-amino:desacetoxy:cephalosporanic acid.
XX
PS  Claim 2; Page; 21pp; English.
XX
XX  AAW34166-W34177 represent mutant expandase sequences of the invention.
CC  These sequences represent mutated versions of the Streptomyces
CC  clavuligerus expandase sequences shown in AAW17775. The DNA encoding the
CC  wild type expandase was subjected to site directed mutagenesis to create
CC  these sequences. The mutant expandase sequences have an altered substrate
CC  specificity compared to the wild type expandase. Expandase catalyses the
CC  expansion of the 5-membered thiazolidine ring of penicillin N to the 6-
CC  membered dihydrothiazine ring of DOAC. The mutant expandase gene is used
CC  particularly for the production of 7-aminodesacetoxycephalosporanic acid
CC  (7-ADCA). 7-ADCA is one of the key intermediates that is used in the
CC  preparation of semi-synthetic cephalosporins. These sequences have
CC  improved ability to expand penicillin G into a cephalosporin ring in
CC  vitro and/or in vivo where other penicillins such as penicillin N and
CC  isopenicillin N can act as competing substrates
XX
SQ  Sequence 311 AA;

Query Match          99.6%; Score 1630; DB 2; Length 311;
Best Local Similarity 99.4%; Pred. No. 2,9e-167;
Matches 309; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFTDLPRPADAVIVECGAIALTLVTCGOYKA 240
DB  181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFTDLPRPADAVIVECGAIALTLVTCGOYKA 240
QY  241 PRHHVAAPRRDQIAGSSSRTSSVFFFLRPNADTFPSVPLARECGFVSLDGETATFQDWIGG 300
DB  241 PRHHVAAPRRDQIAGSSSRTSSVFFFLRPNADTFPSVPLARECGFVSLDGETATFQDWIGG 300
QY  301 NYVNIIRRTSKA 311
DB  301 NYVNIIRRTSKA 311

RESULT 11
AAW34176
ID  AAW34176 standard; protein; 311 AA.
XX
AC  AAW34176;
XX
DT  17-FEB-1998 (first entry)
XX
DE  Expandase mutant R266Q.
XX
KM  Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;
KM  thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;
KM  semi-synthetic cephalosporin production; penicillin G; isopenicillin N;
KM  cephalosporin ring; DOAC; mutant.
XX
OS  Streptomyces clavuligerus.
OS  Synthetic.
XX
FH  Key
FT  Misc-difference 266
FT  Location/Qualifiers
XX  /note= "Arg to Gln mutation"
XX
PN  MO9720053-A2.
XX
PD  05-JUN-1997.
XX
PF  27-NOV-1996; 96WO-EP005358.
XX
PR  27-NOV-1995; 95EP-00203259.
PR  27-NOV-1995; 95US-0007564P.
XX
PA  (KONN ) GIST-BROCADES BV.
XX  (UYOX-) UNIV OXFORD.
XX
PI  Sutherland JD, Bovenberg RAL, Van Der Laan JM;
XX  WPI; 1997-310608/28.
XX
DR  WPI; 1997-310608/28.
XX
PT  Expandase enzymes with mutation(s) to alter substrate specificity - used
PT  particularly for production of 7-amino:desacetoxy:cephalosporanic acid.
XX
PS  Claim 2; Page; 21pp; English.
XX
XX  AAW34166-W34177 represent mutant expandase sequences of the invention.
CC  These sequences represent mutated versions of the Streptomyces
CC  clavuligerus expandase sequences shown in AAW17775. The DNA encoding the
CC  wild type expandase was subjected to site directed mutagenesis to create
CC  these sequences. The mutant expandase sequences have an altered substrate
CC  specificity compared to the wild type expandase. Expandase catalyses the
CC  expansion of the 5-membered thiazolidine ring of penicillin N to the 6-
CC  membered dihydrothiazine ring of DOAC. The mutant expandase gene is used
CC  particularly for the production of 7-aminodesacetoxycephalosporanic acid
CC  (7-ADCA). 7-ADCA is one of the key intermediates that is used in the
CC  preparation of semi-synthetic cephalosporins. These sequences have
CC  improved ability to expand penicillin G into a cephalosporin ring in
CC  vitro and/or in vivo where other penicillins such as penicillin N and
CC  isopenicillin N can act as competing substrates
XX
SQ  Sequence 311 AA;

```

Query Match 99.6%; Score 1630; DB 2; Length 311;
 Best Local Similarity 99.4%; Pred. No. 2.9e-167;
 Matches 309; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTTVPFSLAELOQGLHODEFRRCCLRDKGLFYLTDCGLTDTTELKSAKDIVIDFPEHSGE 60
 Db 1 MDTTVPFSLAELOQGLHODEFRRCCLRDKGLFYLTDCGLTDTTELKSAKDIVIDFPEHSGE 60
 QY 61 AEKRAVTSVPVPTMRKRGFTGLESESTAOITNTGYSYSDYSCYSGMTADNLFPSGDFERIMT 120
 Db 61 AEKRAVTSVPVPTMRKRGFTGLESESTAOITNTGYSYSDYSCYSGMTADNLFPSGDFERIMT 120
 QY 121 QYFDRQYTASRAVAREVLRAATGTEPDGVEAFIDCEPLRFRFPQVPEHRSABEQPLRM 180
 Db 121 QYFDRQYTASRAVAREVLRAATGTEPDGVEAFIDCEPLRFRFPQVPEHRSABEQPLRM 180
 QY 181 APHYDLSMWTLLIQOTPCANGFVSLQAEVGAFTDLPRPDVAVLFCGAIATLVTTGGQVKA 240
 Db 181 APHYDLSMWTLLIQOTPCANGFVSLQAEVGAFTDLPRPDVAVLFCGAIATLVTTGGQVKA 240
 QY 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATFQDWIGG 300
 Db 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATFQDWIGG 300
 QY 301 NYVINIRRTSKA 311
 Db 301 NYVINIRRTSKA 311

RESULT 12

AAW40424
 ID AAW40424 standard; protein; 311 AA.

AC AAW40424;

DT 02-JUL-1998 (first entry)

DE S. clavuligerus expandase protein.

KW Expandase; adipoyl cephalosporin; desacetoxyccephalosporin C synthase;

KM DAOCS; 5-carboxypentanoyl cephalosporin.

OS Streptomycetes clavuligerus.

XX MO9802551-A2.

XX 22-JAN-1998.

XX 15-JUL-1997; 97WO-EP003879.

XX 16-JUL-1996; 96EP-00201988.

XX (KONN) GIST-BROCADES BV.

PI Bovenberg RAL, Van Der laan JM, Kerkman R, Nieboer M;

XX WPI; 1998-110597/10.

PT Modified expandase genes - which encode mutant expandase with amino acid
 substitutions used for the production of cephalosporin compounds.

PS Disclosure; Page; 48pp; English.

XX This sequence represents the expandase protein (also known as
 CC desacetoxyccephalosporin C synthase or DAOCS) from Streptococcus
 CC clavuligerus. Mutants of this protein are used in a novel method
 CC involving a biosynthetic process for preparation and recovery of adipoyl
 CC cephalosporins (5-carboxypentanoyl cephalosporins). The modified
 CC expandase provides for the more efficient production of cephalosporins
 XX Sequence 311 AA;

Query Match 99.6%; Score 1630; DB 2; Length 311;
 Best Local Similarity 99.7%; Pred. No. 2.9e-167;
 Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTVPFSLAELOQGLHODEFRRCCLRDKGLFYLTDCGLTDTTELKSAKDIVIDFPEHSGE 60
 Db 1 MDTTVPFSLAELOQGLHODEFRRCCLRDKGLFYLTDCGLTDTTELKSAKDIVIDFPEHSGE 60
 QY 61 AEKRAVTSVPVPTMRKRGFTGLESESTAOITNTGYSYSDYSCYSGMTADNLFPSGDFERIMT 120
 Db 61 AEKRAVTSVPVPTMRKRGFTGLESESTAOITNTGYSYSDYSCYSGMTADNLFPSGDFERIMT 120
 QY 121 QYFDRQYTASRAVAREVLRAATGTEPDGVEAFIDCEPLRFRFPQVPEHRSABEQPLRM 180
 Db 121 QYFDRQYTASRAVAREVLRAATGTEPDGVEAFIDCEPLRFRFPQVPEHRSABEQPLRM 180
 QY 181 APHYDLSMWTLLIQOTPCANGFVSLQAEVGAFTDLPRPDVAVLFCGAIATLVTTGGQVKA 240
 Db 181 APHYDLSMWTLLIQOTPCANGFVSLQAEVGAFTDLPRPDVAVLFCGAIATLVTTGGQVKA 240
 QY 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATFQDWIGG 300
 Db 241 PRHHVAAPRDQIAGSSRTSSVFFLRPNADFTFSVPLARECGDFVSLDGETATFQDWIGG 300
 QY 301 NYVINIRRTSKA 311
 Db 301 NYVINIRRTSKA 311

RESULT 13

AAW34175
 ID AAW34175 standard; protein; 311 AA.

AC AAW34175;

DT 17-FEB-1998 (first entry)

DE Expandase mutant R266N.

KW Expandase; 7-aminodesacetoxycephalosporanic acid; 7-ADCA production;

KM thiazolidine ring expansion; penicillin N; dihydrothiazine ring; enzyme;

KM semi-synthetic cephalosporin production; penicillin G; isopenicillin N;

XX cephalosporin ring; DOAC; mutant.

XX Streptomycetes clavuligerus.

XX Synthetic.

XX MO9720053-A2.

XX 05-JUN-1997.

XX 27-NOV-1996; 96WO-EP005358.

XX 27-NOV-1995; 95EP-00203259.

XX 27-NOV-1995; 95US-0007564P.

XX (KONN) GIST-BROCADES BV.

PI (UYOX-) UNIV OXFORD.

XX Sutherland JD, Bovenberg RAL, Van Der laan JM;

XX WPI; 1997-310608/28.

PT Expandase enzymes with mutation(s) to alter substrate specificity - used
 particularly for production of 7-amino:desacetoxy:cephalosporanic acid.

PS Claim 2; Page; 21pp; English.

XX AAW34166-W34177 represent mutant expandase sequences of the invention.

CC These sequences represent mutated versions of the Streptomyces
CC clavuligerus expandase sequences shown in AA017775. The DNA encoding the
CC wild type expandase was subjected to site directed mutagenesis to create
CC these sequences. The mutant expandase sequences have an altered substrate
CC specificity compared to the wild type expandase. Expandase catalyses the
CC expansion of the 5-membered thiazolidine ring of penicillin N to the 6-
CC membered dihydrothiazine ring of DOAC. The mutant expandase gene is used
CC particularly for the production of 7-aminodeacetoxycephalosporanic acid
CC (7-ADCA). 7-ADCA is one of the key intermediates that is used in the
CC preparation of semi-synthetic cephalosporins. These sequences have
CC improved ability to expand penicillin G into a cephalosporin ring in
CC vitro and/or in vivo where other penicillins such as penicillin N and
CC isopenicillin N can act as competing substrates

XX SQ Sequence 311 AA;

Query Match 99.6%; Score 1629; DB 2; Length 311;
Best Local Similarity 99.4%; Pred. No. 3.7e-167;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTVPFSLAELOQGLHODEFRRCRLDKGLFYLTDCGLTDTLKSADVIIDFPHGSE 60
DB 1 MDTTVPFSLAELOQGLHODEFRRCRLDKGLFYLTDCGLTDTLKSADVIIDFPHGSE 60
QY 61 AEKRAVTSVPVPMRGRFTGLSESTAOITNTGSDYSMSYMGTAADNLPFGSGFERIWT 120
DB 61 AEKRAVTSVPVPMRGRFTGLSESTAOITNTGSDYSMSYMGTAADNLPFGSGFERIWT 120
QY 121 QYFDRQYTSRAVAAREVLRAATGTEPDGVEAFIDCEPLLRFRRYPQVPEHRSABEQPLRM 180
DB 121 QYFDRQYTSRAVAAREVLRAATGTEPDGVEAFIDCEPLLRFRRYPQVPEHRSABEQPLRM 180
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQYKA 240
DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQYKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFVSLDGETATFODWIG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFVSLDGETATFODWIG 300
QY 301 NYVNIRRTSKA 311
DB 301 NYVNIRRTSKA 311

RESULT 14
AA012016 standard; protein; 311 AA.
XX AAE12016;
AC AAE12016;
XX 18-DEC-2001 (first entry)
DT Streptomyces clavuligerus modified deacetoxycephalosporin C synthetase.
XX Streptomyces clavuligerus modified deacetoxycephalosporin C synthetase.
XX Deacetoxycephalosporin C synthetase; DAOCS; antibiotic; penicillin;
KW oxygenase; non-haeme iron (II) dependent family; oxidase; mutant; muten.
XX Streptomyces clavuligerus.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 181
FT /note= "Wild type Ala substituted with Arg"
XX US6284483-B1.
XX 04-SEP-2001.
XX 06-OCT-1999; 99US-00413231.
XX 06-OCT-1999; 99US-00413231.
XX 06-OCT-1999; 99US-00413231.

PA (UNMS) UNIV MICHIGAN STATE.
XX Dilley DR, Kadyzhnova DK, Wang Z, Warner TM;
PI WPI; 2001-615433/71.
XX Method for producing antibiotics, particularly penicillin G or V,
PT comprises employing a modified isopenicillin N synthetase, in either an
PT organism or a cell-free system under the control of bicarbonate.
XX Disclosure; Col 45-48; 27pp; English.
PS The invention relates to a modified enzyme of a non-haeme iron (II)
CC dependent family of oxygenases and oxidases which renders the enzyme
CC dependent on bicarbonate for activity. The invention also related to a
CC method for producing penicillin G or V comprises employing a modified
CC enzyme, particularly isopenicillin N synthetase (IPNS), in either an
CC organism or a cell-free system, under the control of bicarbonate. The
CC modified enzymes are IPNS, deacetoxycephalosporin C synthetase (DAOCS)
CC and deacetoxycephalosporin C synthetase (DAOCS). The method is useful for
CC producing antibiotics, particularly penicillin G or V. The method is
CC useful for making organisms useful for making an antibiotic dependent on
CC bicarbonate to make the antibiotic. The present sequence is Streptomyces
CC clavuligerus modified deacetoxycephalosporin C synthetase showing an
CC Arg181 in place of Ala181

XX SQ Sequence 311 AA;

Query Match 99.6%; Score 1629; DB 4; Length 311;
Best Local Similarity 99.4%; Pred. No. 3.7e-167;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTTVPFSLAELOQGLHODEFRRCRLDKGLFYLTDCGLTDTLKSADVIIDFPHGSE 60
DB 1 MDTTVPFSLAELOQGLHODEFRRCRLDKGLFYLTDCGLTDTLKSADVIIDFPHGSE 60
QY 61 AEKRAVTSVPVPMRGRFTGLSESTAOITNTGSDYSMSYMGTAADNLPFGSGFERIWT 120
DB 61 AEKRAVTSVPVPMRGRFTGLSESTAOITNTGSDYSMSYMGTAADNLPFGSGFERIWT 120
QY 121 QYFDRQYTSRAVAAREVLRAATGTEPDGVEAFIDCEPLLRFRRYPQVPEHRSABEQPLRM 180
DB 121 QYFDRQYTSRAVAAREVLRAATGTEPDGVEAFIDCEPLLRFRRYPQVPEHRSABEQPLRM 180
QY 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQYKA 240
DB 181 APHYDLSWVTLIQOTPCANGFVSLQAEVGAFITDLPYRPDAVLVFCGAIATLVGGQYKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFVSLDGETATFODWIG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECGFVSLDGETATFODWIG 300
QY 301 NYVNIRRTSKA 311
DB 301 NYVNIRRTSKA 311

RESULT 15
AAU1047 standard; protein; 311 AA.
XX AAU1047;
AC AAU1047;
XX 12-MAR-2002 (first entry)
DT Streptomyces clavuligerus penicillin N expandase L158A mutant.
XX Streptomyces clavuligerus penicillin N expandase L158A mutant.
XX Penicillin N expandase; ring-expanding; penicillin G;
KW phenylacetyl-7-ADCA; amino deacetoxycephalosporanic acid; cephalosporin;
KW cephalaxin; penicillin V; mutant; muten.
XX Streptomyces clavuligerus.
OS Synthetic.

J. Bacteriol. 172, 3952-3958, 1990
A:Title: The beta-lactam biosynthesis genes for isopenicillin N epimerase and deacetoxy
A:Reference number: Z26033; MUID:90299822; PMID:1694525
A:Accession: T52312
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-311 <KOV>
A:Cross-references: UNIPROT:P18548; EMBL:M32324; PIDN:AAA6715.1
C:Genetics:
A:Gene: cefE
C:Superfamily: isopenicillin N synthase

Query Match 99.9%; Score 1634; DB 2; Length 311;
Best Local Similarity 99.7%; Pred. No. 4.7e-139;
Matches 310; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTVPFSLAELOQGLHODEFRRLCDKGLFYLTDGCLTDTELKSAKDIVIDFEHGE 60
DB 1 MDTVPFSLAELOQGLHODEFRRLCDKGLFYLTDGCLTDTELKSAKDIVIDFEHGE 60
QY 61 AEKAAVTSVPVPMRGFTGLESESTAOITNTGSDYSVMCMTADNLPFSGDFERLWT 120
DB 61 AEKAAVTSVPVPMRGFTGLESESTAOITNTGSDYSVMCMTADNLPFSGDFERLWT 120
QY 121 QYFPROYTSARAVAREVLRAATGTEPPDGVAEAFLDCEPLRFRYPPOVPEHRSABEOPLRM 180
DB 121 QYFPROYTSARAVAREVLRAATGTEPPDGVAEAFLDCEPLRFRYPPOVPEHRSABEOPLRM 180
QY 181 APHYDLSMVTLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLFCAIATLVTGQVKA 240
DB 181 APHYDLSMVTLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLFCAIATLVTGQVKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
QY 301 NYVINIRRTSKA 311
DB 301 NYVINIRRTSKA 311

RESULT 3

S30900
deacetoxycephalosporin C synthetase - Streptomyces lactamdurans
C:Species: Streptomyces lactamdurans
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: S30900; S28383
R:Coque, J.J.R.; Martin, J.F.; Liras, P.
Mol. Gen. Genet. 236, 453-458, 1993
A:Title: Characterization and expression in Streptomyces lividans of cefD and cefE genes
clavuligerus.
A:Reference number: S30900; MUID:93173127; PMID:8437592
A:Accession: S30900
A:Molecule type: DNA
A:Residues: 1-314 <COO>
A:Cross-references: UNIPROT:Q03047; EMBL:Z13974; NID:949299; PIDN:CAA78376.1; PID:949300
A:Note: the source is designated as Nocardia lactamdurans
C:Genetics:
A:Gene: cefE
C:Superfamily: isopenicillin N synthase

Query Match 73.1%; Score 1195.5; DB 2; Length 314;
Best Local Similarity 70.4%; Pred. No. 1.2e-99;
Matches 219; Conservative 40; Mismatches 51; Indels 1; Gaps 1;

QY 2 DTTVPFSLAELOQGLHODEFRRLCDKGLFYLTDGCLTDTELKSAKDIVIDFEHGE 60
DB 3 DATVPFDLAEIRGLQGEFRHCLREKGVYLTGTLPAADHAGRELAVIDFHDGTE 62
QY 61 AEKAAVTSVPVPMRGFTGLESESTAOITNTGSDYSVMCMTADNLPFSGDFERLWT 120
DB 63 AEKAAVTSVPVPMRGFTGLESESTAOITNTGSDYSVMCMTADNLPFSGDFERLWT 122

QY 121 QYFPROYTSARAVAREVLRAATGTEPPDGVAEAFLDCEPLRFRYPPOVPEHRSABEOPLRM 180
DB 123 DYFARMTRASQDVARQVLSVGAPEVGNDAFLDCPLRLKRYFPEVPEHRSABEOPLRM 182
QY 181 APHYDLSMVTLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLFCAIATLVTGQVKA 240
DB 183 APHYDLSIVTLIHQTCANGFVSLQAEVGVDSYVDIPAGCAVLVFCGAVATLVADGAIRA 242
QY 241 PRHHVAAPRRDQIAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
DB 243 PRHHVAAPRRDQIAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 302
QY 301 NYVINIRRTSKA 311
DB 303 NYVINIRRTSKA 313

RESULT 4

S40253
deacetoxycephalosporin C synthetase - Streptomyces lactamdurans
C:Species: Streptomyces lactamdurans
C>Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: S40253
R:Coque, J.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40253
A:Accession: S40253
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <COO>
A:Cross-references: UNIPROT:P42219; EMBL:Z21687; NID:9438193; PIDN:CAA79803.1; PID:943819;
C:Note: the source is designated as Nocardia lactamdurans
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 59.4%; Score 972.5; DB 2; Length 310;
Best Local Similarity 61.5%; Pred. No. 1.3e-79;
Matches 187; Conservative 37; Mismatches 79; Indels 1; Gaps 1;

QY 2 DTTVPFSLAELOQGLHODEFRRLCDKGLFYLTDGCLTDTELKSAKDIVIDFEHGE 60
DB 3 DKTVPFSLAELOQGLHODEFRRLCDKGLFYLTDGCLTDTELKSAKDIVIDFEHGE 61
QY 62 EKAATVTSVPVPMRGFTGLESESTAOITNTGSDYSVMCMTADNLPFSGDFERLWTQ 121
DB 62 EKAATVTSVPVPMRGFTGLESESTAOITNTGSDYSVMCMTADNLPFSGDFERLWTQ 121
QY 122 YEDROYTSARAVAREVLRAATGTEPPDGVAEAFLDCEPLRFRYPPOVPEHRSABEOPLRMA 181
DB 122 YEDROYTSARAVAREVLRAATGTEPPDGVAEAFLDCEPLRFRYPPOVPEHRSABEOPLRMA 181
QY 182 PHYDLSMVTLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLFCAIATLVTGQVKA 241
DB 182 PHYDLSMVTLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLFCAIATLVTGQVKA 241
QY 242 RHHVAAPRRDQIAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 301
DB 242 RHHVAAPRRDQIAGSSRTSSVFELRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 301
QY 302 YVNI 305
DB 302 YVNI 305

RESULT 5

A39204
deacetoxycephalosporin C synthetase hydroxylase - Streptomyces clavuligerus
C:Species: Streptomyces clavuligerus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A39204; A23713
R:Kovacevic, S.; Miller, J.R.
J. Bacteriol. 173, 398-400, 1991
A:Title: Cloning and sequencing of the beta-lactam hydroxylase gene (cefE) from Streptomyces

A:Reference number: A39204; MUID:91100311, PMID:1987130
A:Accession: A39204
A:Molecule type: DNA
A:Residues: 1-318 <KOV>
A:Cross-references: UNIPROT:P42220; GB:M63809; GB:M37186; NID:g153206; PIDN:AAA26716.1;
R:Baker, B.J.; Dotzaf, J.E.; Yeh, W.K.
J. Biol. Chem. 266, 5087-5093, 1991
A>Title: Desacetoxycephalosporin C hydroxylase of Streptomyces clavuligenus. Purification
A:Reference number: A33713; MUID:91161600; PMID:2002049
A:Accession: A33713
A:Molecule type: protein
A:Residues: 2-29/32-100 <BAK>
C:Genetics:
A:Gene: cefP
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 57.9%; Score 946.5; DB 2; Length 318;
Best Local Similarity 59.0%; Pred. No. 2.8e-77;
Matches 181; Conservative 43; Mismatches 80; Indels 3; Gaps 1;

Dy DTTVPFSLAELOOGLHODEFRRCIRDKGLFYLTDCGTLDTLKSANDIVDFEHGSEA 61
DTPVPIFIALLAREGADKEKRECVTGMGVFLYLGAGDGKHRLATDTAMDPAFGCTEA 62
EKAAYTSVPMRMGRFTGLSESTAOITNTSYSDYSIMCGTADNLFPBGDEPERIWTQ 121
EKAAYTVDPFMRRGSYSLAEASTAQVTRTSYTDSFSMGISGNVFSPPEPERVWTE 122
YFDROYTSRAVARREVLRATG--TERPDGYEAFIDCEPLIRFRYPPOVPRHSABROPL 178
YFDKLIAAAQETARLVLTASGGYDAELVIGSIDELLIDAPVLRILRFEPVPRHSABRPR 182
RMAHYDISMTLLIQOTPCANGFVSLQAEVGAFETDLYPRDAVLVFCGATATLVTSQOV 238
RMAHYDLSTITTFHQPTCANGFVSLQAEICGEVLSLPVEDAVVMCGAAMALATGAL 242
KAPRHVAAPRRDIAGSRSSVFFLRPNADFRTFSPVLAECGFDSLDGETITFDQWI 298
PAPRHHRNSPGAKMRSGSDRTSVFFLRPTDFSFSAKARSYGIALVDLMETATFDGI 302

Oy 299 GGNYVNI 305
| | | |
Db 303 GTNVVTM 309

RESULT 6
SS4100
desacetoxycephalosporin C synthetase 1 - lysobacter lactamgens (strain YK90)
C:Species: Lysobacter lactamgens
A:Variety: strain YK90
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: SS4100
R:Kimura, H.; Izawa, M.; Miyashita, H.; Shimizu, Y.; Sumino, Y.; Suzuki, M.
submitted to the EMBL Data Library, October 1990
A:Description: Gene cluster involved in the cephalosporin biosynthesis from Lysobacter 1
A:Reference number: SS4099
A:Accession: SS4100
A:Molecule type: DNA
A:Residues: 1-319 <KIT>
A:Cross-references: UNIPROT:Q48740; EMBL:X56660; NID:g769806; PIDN:CAA39984.1; PTD:g769806
A:Experimental source: strain YK90
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase
C:Keywords: antibiotic biosynthesis

Query Match 57.7%; Score 944; DB 2; Length 319;
Best Local Similarity 58.4%; Pred. No. 4.8e-77;
Matches 178; Conservative 45; Mismatches 82; Indels 0; Gaps 0;

Oy DTTVPFSLAELOOGLHODEFRRCIRDKGLFYLTDCGTLDTLKSANDIVDFEHGSEA 61
DTPVPIFIALLAREGADKEKRECVTGMGVFLYLGAGDGKHRLATDTAMDPAFGCTEA 62
DSGQIPIPLDLERHGVRNDRSKSLFERGCVYVAREDSIKTEHAKKAMDAVMDLFENGSAE 62
EKRAVTSVPMRMGRFTGLSESTAOITNTSYSDYSIMCGTADNLFPBGDEPERIWTQ 121

[illegible]

RESULT 11

S08218

isopenicillin N synthase (EC 1.14.11.-) pcbC [similarity] - Flavobacterium sp. (strain S08218)

C:Species: Flavobacterium sp.

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 26-May-2000

C:Accession: S08218

R:Shiffman, D.; Cohen, G.; Aharonowitz, Y.; Palissa, H.; von Dohren, H.; Kleinkauf, H.;

Nucleic Acids Res. 18, 660, 1990

A:Title: Nucleotide sequence of the isopenicillin N synthase gene (pcbC) of the gram neg

A:Reference number: S08218; MUID:90174998; PMID:2308852

A:Accession: S08218

A:Molecule type: DNA

A:Residues: 1-326 <SHI>

A:Cross-references: EMBL:X17355; NID:g43423; PIDN:CAA35233.1; PID:g43424

A:Gene: pcbC

C:Gene: pcbC

C:Superfamily: isopenicillin N synthase

C:Keywords: antibiotic biosynthesis; ascorbic acid; iron; metalloprotein; oxidoreductase

F:45,209,265/Binding site: iron (His) #status predicted

Query Match

Best Local Similarity 10.0%; Score 164; DB 2; Length 326;

Matches 65; Conservative 34; Mismatches 107; Indels 78; Gaps 11;

QY

24 RCLRDKGLFYLTDCGLTDELKSAKDVIYDFEHGSEAK-----RAVTSVPVTRRGFT 78

DB

32 RACRSGGFYAAHNGV--DLAALQKFTTDMHMANSAEKWELAIRAVNPANPRNRNGY- 87

QY

79 GLESESTAQITNTGYSYDMCSMGCTAD-----NLFPS---GDFERIW 119

DB

88 -----YMAVEGKANNSFCYLNPSFDADHATIKAGLPSHEVINIWDEARHPGMRFFY 139

QY

120 TOYFDROYTASRAVAEVLRAATGTEPDGVEAF-----LDCEPLLRFRYPQVPE 169

DB

140 EAYFSDVDVAAVILIRGFAIALGRE-----ESFERRHFSMDTILSAVSLIRYPLENYP- 193

QY

170 HRSABEQLRMAP-----HYDLSWTLIQOTPCANGFVSLQAEVGAGATDLPYRPD 220

DB

194 -----PLKIGPDEKLSFEHHQDVSLITLYQTAPN-----LQVETAEGLDIPVSD 242

QY

221 AVLVFCGALATLVGGOVKAPRHVAAPRRQIAGSRTSSVFF 264

DB

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

DB

32 RACRSGGFYAAHNGV--DLAALQKFTTDMHMANSPBEKWELAIRAVNPANPRNRNGY- 87

QY

79 GLESESTAQITNTGYSYDMCSMGCTAD-----NLFPS---GDFERIW 119

DB

88 -----YMAVEGKANNSFCYLNPSFDADHATIKAGLPSHEVINIWDEARHPGMRFFY 139

QY

120 TOYFDROYTASRAVAEVLRAATGTEPDGVEAF-----LDCEPLLRFRYPQVPE 169

DB

140 EAYFSDVDVAAVILIRGFAIALGRE-----ESFERRHFSMDTILSAVSLIRYPLENYP- 193

QY

170 HRSABEQLRMAP-----HYDLSWTLIQOTPCANGFVSLQAEVGAGATDLPYRPD 220

DB

194 -----PLKIGPDEKLSFEHHQDVSLITLYQTAPN-----LQVETAEGLDIPVSD 242

QY

221 AVLVFCGALATLVGGOVKAPRHVAAPRRQIAGSRTSSVFF 264

DB

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

243 HFLVNCGYMAHTNGYYPAPVHRV-----KYNAERLSIFPF 280

QY

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C64713
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STO>
A:Cross-references: UNIPROT:O80850; GB:AE002093; NID:g3201632; PIDN:AA020718.1; GSPDB:GN
C:Genetics:
A:Gene: At2g30830
A:Map position: 2
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 8.8%; Score 144.5; DB 2; Length 358;
Best Local Similarity 22.3%; Pred. No. 3.7e-05;
Matches 60; Conservative 47; Mismatches 116; Indels 41; Gaps 9;

QY 23 RCLRDKGLFYLTDCGLTDTLTKSAKDVIDFFHSGSEAKRAVTSPPVPMRGRFTGLSES 82
DB 78 RAAVEKEGFFQVINHGIPLEVESMKDGIKGFHEODSEVKTFYSRDI-----T 126
QY 83 ESTAQITNTGSDYSIDSMCY---SMGTADNLFPSGDF---ERIWTQYFDRQYASRAVA 134
DB 127 KKVKYNTNFDIYSSQAAMWRDTLTWMAAPDVPAQADLPVICREIMLEYSKMMKLGELIF 186
QY 135 REVLARGTGTEPDGGEAFLEDCPELIRF--RYFPQVPEHRSABEQPLRNAPHYDLSWYTLI 192
DB 187 ELTSEALGLKPNHKE--LNCAKSLSLSHYPPCPD---PDRFGISSHTDISFITIL 240
QY 193 QQTCCANGFVSLQAEVGAFTDLPYRPDAVLFCGAIATLTVGGQVAKAPRHVAAPRRDQ 252
DB 241 LQ---DHIGGLQVLHDGYWIDPVPNPALIVNLGDLQLITNDKQVSEHRAVLNRGE 296
QY 253 IAGSSRTSSVFPLRNADFTFSVP 276
DB 297 ----PRISSASF-----FMHTIP 310

RESULT 15
S33510
flavonol synthase - garden petunia
C:Species: Petunia x hybrida (garden petunia)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S33510
R:Holt, T.A.
submitted to the EMBL Data Library, April 1993
A:Description: Cloning and expression of flavonol synthase from Petunia hybrida.
A:Reference number: S33510
A:Accession: S33510
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-348 <HOL>
A:Cross-references: UNIPROT:Q07512; EMBL:Z22543; NID:g311657; PIDN:CA80264.1; PID:g3116
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase
C:Keywords: iron; metalloprotein
F:87,234,290/Binding site: Iron (His) #status predicted

Query Match 8.7%; Score 143; DB 2; Length 348;
Best Local Similarity 22.3%; Pred. No. 4.9e-05;
Matches 59; Conservative 43; Mismatches 114; Indels 48; Gaps 9;

QY 27 RDKGLFYLTDCGLTDTLTKSAKDIYDFFHSGSEAKRAVT-SPVPTMRGRF-TGLSESES 84
DB 77 KEMGIFQILNHGIDPEADLQKVKEFEHVPQBEKELIAKTPGSNDIBGYSTLQKEV 136
QY 85 TAOITNTGSDYSIDSMCYSMGTADNLF-----PSGDFERIW-----TOYFDRQYT 128
DB 137 EGR-----KGVNDHLFKIKIPPSAVNRYWPKNPSPSYREANEELYGKMRRE 181
QY 129 ASRAVAREVLARGTGTEPDGGEAFLEDC--PLRFRYPQVPEHRSABEQPLRNAPHYDL 186

DB 182 VVDRIFKSLSLGLGHEHMEIAAGDEIVLLKINYPPCPR----PDALGVVAHTMD 237
QY 187 SWYTLIQQTPCANGFVSLQAEVGAFTDLPYRPDAVLFCGAIATLVTGGQVAKAPRHVA 246
DB 238 SYITIL-----VPEVVOGLQVFKDGHWDVXYIPNALIVHIGQVEILSNKYSVYHRTT 293
QY 247 APRRDQIAGSSRTSSVFPLRNAD 270
DB 294 VNK-----DKTRMSWVFLEPPSE 312

Search completed: September 19, 2005, 15:18:50
Job time : 32 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein, Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 19, 2005, 15:01:52; Search time 90 Seconds

(without alignments)
1769.518 Million cell updates/sec

Title: 10719236-1_232-1164

Perfect score: 1636
Sequence: 1 MDTWPTFSLAEIQGLHQD.....ATFDWIGGNVYNIIRTSKA 311

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_prot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1634	99.9	311	1	CEFE_STRCL
2	1385	84.7	311	1	Q7MYW9
3	1378	84.2	311	2	Q9JFD4
4	1300	79.5	311	2	Q7MYW8
5	1195.5	73.1	314	1	CEFE_NOCUA
6	1012	61.9	312	2	Q7MYW7
7	972.5	59.4	310	1	CEFE_NOCUA
8	946.5	57.9	318	1	CEFE_STRCL
9	944	57.7	319	2	Q48740
10	933.5	57.1	332	1	EXSA_CEPAC
11	924.5	55.5	332	2	Q9P4T5
12	859	52.5	313	2	Q48741
13	427	26.1	82	2	Q7B3R2
14	183	11.2	329	1	IPNS_STRUJ
15	180	11.0	327	2	Q6WGL3
16	178	10.9	329	1	IPNS_STRGR
17	169	10.3	366	2	Q987V6
18	168	10.3	329	1	IPNS_STRCL
19	168	10.3	348	2	Q6N7A0
20	166.5	10.2	337	2	Q881B6
21	165	10.1	329	2	Q9ANU0
22	164	10.0	326	1	IPNS_FLASS
23	164	10.0	326	2	IPNS_LYSIA
24	163	10.0	262	2	Q9PAC2
25	161	9.8	342	2	Q83YD9
26	160	9.8	262	2	Q9PAC0
27	158	9.7	262	2	Q9PAB6
28	158	9.7	262	2	Q9PAC1
29	156.5	9.6	403	2	Q8WZG6
30	156.5	9.6	403	2	Q9VCZ2
31	156	9.5	262	2	Q9PAC3

32	155.5	9.5	408	2	Q614Y3	Q614Y3 oryza sativ
33	155	9.5	262	2	Q9FAB8	Q9fab8 streptomyc
34	155	9.5	262	2	Q9FAB9	Q9fab9 streptomyc
35	155	9.5	331	2	Q9LCZ2	Q9lc22 streptomyc
36	154.5	9.4	366	2	Q942Y0	Q942y0 oryza sativ
37	152	9.3	321	2	Q889J1	Q889j1 pseudomonas
38	147	9.0	333	1	IPNS_STRUP	IP12438 streptomyc
39	146.5	9.0	348	2	Q8SB89	Q8sb89 oryza sativ
40	146.5	9.0	348	2	Q7XC81	Q7xc81 oryza sativ
41	145	8.9	362	2	Q751P9	Q751p9 oryza sativ
42	145	8.9	367	2	Q761Z9	Q761z9 ipomoea nil
43	144.5	8.8	358	2	Q808S0	Q808s0 arabidopsis
44	144.5	8.8	389	2	Q8RVF5	Q8rvf5 oryza sativ
45	144.5	8.8	389	2	Q7EC22	Q7ec22 oryza sativ

ALIGNMENTS

```
RESULT 1
CEFE_STRCL STANDARD, PRT; 311 AA.
ID CEFE_STRCL
AC P18548;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DROCS)
DE (Expandase).
GN Name=cefe;
OS Streptomyces clavuligerus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1901;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;
RX MEDLINE=89123150; PubMed=2644235;
RA Kovacevic S., Weigel B.J., Tobin M.B., Ingolia T.D., Miller J.R.;
RT "Cloning, characterization, and expression in Escherichia coli of the
RT Streptomyces clavuligerus gene encoding deacetoxycephalosporin C
RT synthetase."
RL J. Bacteriol. 171:754-760(1989).
CC -! FUNCTION: Catalyzes the step from penicillin N to deacetoxy-
CC cephalosporin C.
CC -! CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =
CC -! COFACTOR: Iron and ascorbate.
CC -! PATHWAY: Cephalosporin antibiotics biosynthesis.
CC -! SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase
CC family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M32324; AAA26715.1; -.
CC PIR: T52312; T52312.
CC PDB: 1DCS; X-ray; A=1-311.
CC PDB: 1ESU; X-ray; A=1-306.
CC PDB: 1ES1; X-ray; A=1-306.
CC PDB: 1HUJ; X-ray; A=1-311.
CC PDB: 1HUJ; X-ray; A=1-311.
CC PDB: 1RXF; X-ray; @=1-311.
CC PDB: 1RXG; X-ray; @=1-311.
CC PDB: 1UNB; X-ray; A=1-311.
CC PDB: 1UO9; X-ray; A=1-311.
CC PDB: 1UO8; X-ray; A=1-311.
CC PDB: 1UOF; X-ray; A=1-311.
CC PDB: 1UOG; X-ray; A=1-311.
```

```
DR InterPro: IPR005123; 2OG-Fer1_Oxy.
DR InterPro: IPR002057; Isopen_N_synth.
DR Pfam: PF03171; 2OG-Fer1_Oxy; 1.
DR PROSITE; PS00185; IPNS_1; FALSE_NEG.
DR PROSITE; PS00186; IPNS_2; 1.
KM 3D-structure; Antibiotic biosynthesis; Iron; Oxidoreductase;
KM Vitamin C.
FT STRAND 7 9
FT HELIX 10 14
FT TURN 15 16
FT HELIX 19 28
FT TURN 29 29
FT STRAND 31 35
FT HELIX 41 57
FT HELIX 60 65
FT TURN 66 66
FT STRAND 67 67
FT STRAND 76 79
FT STRAND 99 103
FT HELIX 113 140
FT TURN 141 142
FT TURN 146 147
FT HELIX 149 153
FT STRAND 158 164
FT STRAND 180 185
FT STRAND 189 195
FT TURN 198 199
FT STRAND 204 208
FT TURN 209 210
FT STRAND 211 214
FT TURN 219 220
FT STRAND 222 226
FT HELIX 228 233
FT TURN 234 235
FT STRAND 243 245
FT TURN 254 255
FT STRAND 258 265
FT TURN 269 270
FT STRAND 272 274
FT HELIX 275 280
FT TURN 281 282
FT STRAND 291 293
FT HELIX 294 298
FT STRAND 305 306
SQ SEQUENCE 311 AA; 34555 MW; 9C64E1FC37F524BC CRC64;

Query Match 99.9%; Score 1634; DB 1; Length 311;
Best Local Similarity 99.7%; Pred. No. 1.5e-139;
Matches 310; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTVPFSLAELOOGLHODEFRRCRLDKGLFYLTDGSLTDTLKSAXDIVIDFPEHGE 60
DB 1 MDTVPFSLAELOOGLHODEFRRCRLDKGLFYLTDGSLTDTLKSAXDIVIDFPEHGE 60
QY AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
DB AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
QY AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
DB AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
QY 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
DB 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
QY 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
DB 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
QY 181 APHYDLSMWTLIQOTPCANGFVSLQAEVGAFTDLPRPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMWTLIQOTPCANGFVSLQAEVGAFTDLPRPDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDOIGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDOIGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311
```

```
RESULT 2
ID Q7MYW9 PRELIMINARY; PRT; 311 AA.
AC Q7MYW9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Deacetoxycephalosporin C synthase.
OS Streptomyces ambolicus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1899;
RN [1]
RP SEQUENCE FROM N.A.
RA Heu J.S., Yang Y.B., Wei C.L., Tsai Y.C.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase
family.
DR EMBL; AY18742; AAP85380.1; -.
DR HSSP; P18548; IE51.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR InterPro; IPR005123; 2OG-Fer1_Oxy.
DR InterPro; IPR002057; Isopen_N_synth.
DR Pfam; PF03171; 2OG-Fer1_Oxy; 1.
DR PROSITE; PS00186; IPNS_2; UNKNOWN_1.
KM Iron; Oxidoreductase; Porin.
SQ SEQUENCE 311 AA; 34396 MW; 190869D7124B96C4 CRC64;

Query Match 84.7%; Score 1385; DB 2; Length 311;
Best Local Similarity 81.4%; Pred. No. 5.5e-117;
Matches 253; Conservative 33; Mismatches 25; Indels 0; Gaps 0;

QY 1 MDTVPFSLAELOOGLHODEFRRCRLDKGLFYLTDGSLTDTLKSAXDIVIDFPEHGE 60
DB 1 MDTVPFSLAELOOGLHODEFRRCRLDKGLFYLTDGSLTDTLKSAXDIVIDFPEHGE 60
QY AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
DB AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
QY AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
DB AEKRAVTSVPVPTMRGFTGLSESTAOITNTGYSYDYSMCSMGTDNLFPSCGDFER1WT 120
QY 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
DB 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
QY 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
DB 121 QYFDROYTASRAVAEVLATGTEPDGVEAFLDCEPLLRFRYPPOVPEHRSABEQPLRM 180
QY 181 APHYDLSMWTLIQOTPCANGFVSLQAEVGAFTDLPRPDAVLVFCGAIATLVGGQVKA 240
DB 181 APHYDLSMWTLIQOTPCANGFVSLQAEVGAFTDLPRPDAVLVFCGAIATLVGGQVKA 240
QY 241 PRHVAAPRRDOIGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
DB 241 PRHVAAPRRDOIGSSRTSSVFFLRPNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
QY 301 NYVNIIRRTSKA 311
DB 301 NYVNIIRRTSKA 311

RESULT 3
ID Q93FD4 PRELIMINARY; PRT; 311 AA.
AC Q93FD4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Deacetoxycephalosporin C synthase.
OS Streptomyces jumonjensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
```

```

OX NCBI_TaxId=1945;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29864;
RA Sim T.S., Sim J.;
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
CC -! SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase
   family.
DR EMBL: AF317908; AAL09460.1; -.
DR HSSP: P18548; IEST.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019867; C:outer membrane; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0015288; F:porin activity; IEA.
DR InterPro: IPR005123; 2OG-FeII_Oxy.
DR Pfam: PF03171; 2OG-FeII_Oxy; 1.
DR PROSITE: PS00186; IPNS 2; UNKNOWN_1.
KW Iron; Oxidoreductase; Porin.
SQ: SEQUENCE 311 AA; 34242 MW; B821ED796A4F5C1A CRC64;

Query Match      84.2%; Score 1378; DB 2; Length 311;
Best Local Similarity 81.7%; Pred. No. 2,4e-116;
Matches 254; Conservative 28; Mismatches 29; Indels 0; Gaps 0;

QY 1 MDTVPPTSLAELOQGLHODEFRRLDKGLFYLTDCGLTDELKSNADIVDFPEHSGE 60
DB 1 MDTVPPTSLAELOQGLHODEFRRLDKGLFYLTDCGLTDELKSNADIVDFPEHSGE 60
QY 61 AEKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
DB 61 AEKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
QY 61 EKKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
DB 61 EKKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
QY 121 QYFDRQYASRAVAEVLRTATGTEPDGVEAFDLCEPLLRFRYPQVPEHRSABEQPLRM 180
DB 121 QYFDRQYASRAVAEVLRTATGTEPDGVEAFDLCEPLLRFRYPQVPEHRSABEQPLRM 180
QY 121 HYFGHMYDVTQGVARQVLKTTGTETEPGVGESFVDCPEPLLRFRYPVEPEDRSABEQPLRM 180
DB 121 HYFGHMYDVTQGVARQVLKTTGTETEPGVGESFVDCPEPLLRFRYPVEPEDRSABEQPLRM 180
QY 181 APHYDLMSWTLTLOQTPCNGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVLTGGQYKA 240
DB 181 APHYDLMSWTLTLOQTPCNGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVLTGGQYKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFVSLDGETATPDQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFVSLDGETATPDQDWIGG 300
QY 301 NYVINIRRTSKA 311
DB 301 NYVINIRRTSKA 311
DB 301 NYVINIRRTSEA 311

RESULT 4
Q7WYW8 PRELIMINARY; PRT; 311 AA.
ID Q7WYW8;
AC Q7WYW8;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Deacetoxycephalosporin C synthase.
OS Streptomyces chartreusii.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxId=1969;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsu J.S., Yang Y.B., Wei C.L., Tsai Y.C.;
RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
CC -! SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase
   family.
DR EMBL: AY318743; AAP85381.1; -.
DR HSSP: P18548; IEST.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019867; C:outer membrane; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.

```

```

DR GO: GO:0015288; F:porin activity; IEA.
DR InterPro: IPR005123; 2OG-FeII_Oxy.
DR InterPro: IPR002057; Isopen_N_synth.
DR Pfam: PF03171; 2OG-FeII_Oxy; 1.
DR PROSITE: PS00186; IPNS 2; UNKNOWN_1.
KW Iron; Oxidoreductase; Porin.
SQ: SEQUENCE 311 AA; 34076 MW; C26880E90221B441 CRC64;

Query Match      79.5%; Score 1300; DB 2; Length 311;
Best Local Similarity 76.5%; Pred. No. 2,8e-109;
Matches 238; Conservative 32; Mismatches 41; Indels 0; Gaps 0;

QY 1 MDTVPPTSLAELOQGLHODEFRRLDKGLFYLTDCGLTDELKSNADIVDFPEHSGE 60
DB 1 MDTVPPTSLAELOQGLHODEFRRLDKGLFYLTDCGLTDELKSNADIVDFPEHSGE 60
QY 61 AEKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
DB 61 AEKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
QY 61 EKKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
DB 61 EKKRAVTSVPVPMRGFTGLSESTAOITNTGYSYDSYCMYSGMTADNLFPSCGFERLWT 120
QY 121 QYFDRQYASRAVAEVLRTATGTEPDGVEAFDLCEPLLRFRYPQVPEHRSABEQPLRM 180
DB 121 QYFDRQYASRAVAEVLRTATGTEPDGVEAFDLCEPLLRFRYPQVPEHRSABEQPLRM 180
QY 121 GYFDRMYDASREVARQVLKTTGTETEPGVGESFVDCPEPLLRFRYPVEPEDRSABEQPLRM 180
DB 121 GYFDRMYDASREVARQVLKTTGTETEPGVGESFVDCPEPLLRFRYPVEPEDRSABEQPLRM 180
QY 181 APHYDLMSWTLTLOQTPCNGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVLTGGQYKA 240
DB 181 APHYDLMSWTLTLOQTPCNGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVLTGGQYKA 240
QY 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFVSLDGETATPDQDWIGG 300
DB 241 PRHHVAAPRRDQIAGSSRTSSVFFLRPNADFTFSPVLARECGFVSLDGETATPDQDWIGG 300
QY 301 NYVINIRRTSKA 311
DB 301 NYVINIRRTSKA 311
DB 301 NYVINIRRTSKA 311

RESULT 5
CEFE_NOCCLA STANDARD; PRT; 314 AA.
ID CEFE_NOCCLA
AC Q03047;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Deacetoxycephalosporin C synthetase (EC 1.14.20.1) (DAOS)
DB Deacetoxycephalosporin C synthetase.
OS Name=cefe;
OS Nocardia lactamdurans.
OC Bacteria; Actinobacteridae; Actinobacteriales;
OC Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
CX NCBI_TaxId=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93173127; PubMed=8437592;
RA Coque J.F.R., Martin J.F., Litras P.;
RT "Characterization and expression in Streptomyces lividans of cefd and
   cefb genes from Nocardia lactamdurans: the organization of the
   cephalosporin gene cluster differs from that in Streptomyces
   clavuligerus."
RL Mol. Gen. Genet. 236:453-458 (1993).
CC -! FUNCTION: Catalyzes the step from penicillin N to deacetoxy-
   cephalosporin C.
CC -! CATALYTIC ACTIVITY: Penicillin N + 2-oxoglutarate + O(2) =
   deacetoxycephalosporin C + succinate + CO(2) + H(2)O.
CC -! COFACTOR: Iron and ascorbate.
CC -! PATHWAY: Cephalosporin antibiotics biosynthesis.
CC -! SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase
   family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC	modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.cn/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; Z13974; CAA78376.1; -
DR	PIR; S30900; S30900.
DR	HSSP; P18548; 1851.
DR	InterPro; IPR005123; 2OG-FelI_Oxy.
DR	InterPro; IPR002057; Isogen_N_synth.
DR	Pfam; PF01171; 2OG-FelI_Oxy; 1.
DR	PROSITE; PS00185; IPNS_1; FALSE NEG.
DR	PROSITE; PS00186; IPNS_2; FALSE NEG.
KW	Antibiotic biosynthesis; Iron; Oxidoreductase; Vitamin C.
SEQUENCE	314 AA; 34530 MW; D007AD3D67A1BDF CRC64;
Query Match 73.1%; Score 1195.5; DB 1; Length 314;	
Best Local Similarity 70.4%; Pred. No. 8.3e-100;	
Matches 219; Conservative 40; Mismatches 51; Indels 1; Gaps 14	
QY	2 DTVVPTFLAEIHOQIHODEFRCRDGLFLITDCG-TPELTSAKDIVIDPFHSG 60
Db	3 DATPTFLAEIRBEIHOEFHRCIREKGVFLYKKTGLPAEDMHSGRGRIANDPFDHTE 62
QY	61 AEKRAVTSVPVTRRRGTGLSESTAOITNTGSYSYDYMCSMGTDNLFPSGDFER1WT 120
Db	63 AEKRAVMPPIPIRIRGVAGLSESTAOITNTGKYDYSMSYMGTDNLFPSEAEFEKAME 122
QY	121 QYPRQVYASPAVAEVEVLRATGTEPDGVVAFIDCEPIILRPYPPOVEHNSAEFQPLRM 180
Db	123 DYFARMYASQVNAQVLITSGAEVEGVGNDAPLDCDEPIRLRYFPEVEDRYAEQPLRM 182
QY	181 APHYDLNMTLLIQTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATLVTTGGQVKA 240
Db	183 APHYDLNMTLLIQTPCANGFVSLQAEVGVGNDAPLDCDEPIRLRYFPEVEDRYAEQPLRM 242
QY	241 PRHVAARRRQIAGSSRSTSVFELRPADFTFSPIARREGCFDVSIDGERTATFDWTIG 300
Db	243 PKHVAAGCAKRRGVSSSRSTSVFELRPNGDFEFSVPARREGCFDVSIPAEATATFDWTIG 302
QY	301 NYVNIARTSKA 311
Db	303 NYINIRKTA 313
RESULT 6	
ID	Q7WYW7 PRELIMINARY; PRT; 312 AA.
AC	Q7WYW7;
DT	01-OCT-2003 (TREMBLrel. 25, Created)
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	Deacetylcephalosporin C synthase.
OS	Streptomyces sp. 65PH1.
OC	Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;
OC	Streptomycinae; Streptomycetaceae; Streptomyces.
OX	NCBI_TaxID=237545;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=65PH1;
RA	Hsu J.S., Yang Y.B., Wei C.L., Tsai Y.C.;
RL	Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
CC	-1- SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase family.
DR	EMBL; AY318744; AAP85382.1; -
DR	HSSP; P18548; 1851.
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0019867; C:outer membrane; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0015288; F:porin activity; IEA.
DR	InterPro; IPR005123; 2OG-FelI_Oxy.
DR	Pfam; PF03171; 2OG-FelI_Oxy; 1.
KW	Iron; Oxidoreductase; Porin.
SEQUENCE	312 AA; 34332 MW; 428DPTA15BB02487 CRC64;

Query Match	61.9%	Score 1012;	DB 2;	Length 312;
Best Local Similarity	60.3%;	Pred. No. 3.5e-83;		
Matches 187;	Conservative 47;	Mismatches 76;	Indels 0;	Gaps 0;
OY	2	DTVPTFSLAELOOGHLDQDEFRRLCKRGLFYLTDLCCQLTPELTKSAKDIVIDFEHSGSEA	61	
		: : : : : :		
Dd	3	DTVPISLSLELRKGADQEKRTCTVMGPFYLTYDVRREBDHQLAADVAMDFEHQTYPE	62	
OY	62	EKRAVTSPPVTMRRGFTGLESESTAQITNTGSYSIDSMCSMGTADNLFPSSGDPERIWTQ	121	
		: : : : : :		
Dd	63	EKEAVTTIKIPMRGRGVALHEASTAQTNTGYTDYSMSMGVSNLPFOREREKWTD	122	
OY	122	YFDQAQYASRAVARREVLRATGEEDPDGVFAFLDCEPLLRRPYPQVEHSHSAEQPLRMA	181	
		: : : : : :		
Dd	123	HFDVLUSARAAVAVINAGTTIDAQMDDLDDCPILRIIRTPVEVEHHRAAEHPRRMA	182	
OY	182	PHYDLSMWTLTIQCPCPCNGEVSLOAEVGAFITDLYPRPDVAVLFECGAIALTVTGGOVKAP	241	
		: : : : : :		
Dd	183	PHYDLSITILLHGRCPCANGFPGIAELIGTMVSLPAVRDAVIVLCGAIAPLATGCAVPAP	242	
OY	242	RHVVAAPRRDIIAGSSTRSSVFIFLRPADTFESVPLARECGFVDSLDEGFATTFODWICGN	301	
		: : : : : : :		
Dd	243	QHNRAPACGLRESDSRMSSVFIFLRPSTDPAFSVPKARTGLDVSDLAETATFGEMWIGTN	302	
OY	302	VYNIRRTSKA 311		
		: : : :		
Dd	303	YTHMATPES 312		
RESULT 7				
CDEF_NOCCLA				
ID	_CDEF_NOCCLA	STANDARD;	PRT;	310 AA.
AC	P42219;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Deacetoxycephalosporin C hydroxylase (SC 1.14.11.-)			
DE	(Deacetylcephalosporin C synthetase) (DACS) (Beta-lactam hydroxylase).			
CN	Name=Cefr,			
OS	Nocardia lactamdurans.			
OC	Bacteria; Actinomycetes; Actinobacteridae; Actinomycetales;			
OC	Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.			
RN	NCBI_Taxid=1913;			
	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96287477; PubMed=8703431;			
RA	Coque J., Enguita F.J., Cardoza R.E., Martin J.F., Liras P.;			
RT	"Characterization of the cefr gene of Nocardia lactamdurans encoding a			
RT	3'-methylcephem hydroxylase different from the 7'-cephem hydroxylase."			
RL	Appl. Microbiol. Biotechnol. 44:605-609(1996).			
CC	-1- FUNCTION: Hydroxylation of deacetoxycephalosporin C in 3' position			
CC	to form deacetylcephalosporin C.			
CC	-1- PATHWAY: Cephalosporin biosynthesis.			
CC	-1- SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase			
CC	family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; Z21867; CAAT79803.1; -.			
DR	PIR; S40253; S40253.			
DR	HSSP; P18548; IRLF.			
DR	InterPro; IPR005123; 2OG-Fell_Oxy.			
DR	Pfam; PF03107; 2OG-Fell_Oxy.I.			
KM	Antibiotic biosynthesis; Iron; Oxidoreductase; Vitamin C.			
SQ	SEQUENCE 310 AA; 34364 MW; 99835FFIAC1B416B6 CRC64;			

Query Match 59.4%; Score 972.5; DB 1; Length 310;
 Best Local Similarity 61.5%; Pred. No. 1.3e-79;
 Matches 187; Conservative 37; Mismatches 79; Indels 1; Gaps 1;

QY 2 DTTVPTFSLAEIQQGLHODEFRRLCKDKLFYLTDCGLTDTLTKSAKDIVIDPFHSGSEA 61
 DB 3 DTPVPIFNLALREGADDEKRECVTGWGVYLTGAGDKDHLRTATDTAMDFANGTEA 62
 QY 62 EKRAVTSVPVPMRGRGFTGLSESTAOITNTGSSYSDSYCMYSGMGTADNLPFGSDPERITWQ 121
 DB 62 EKAAVTTDVPMRGRGYSALAEASTAOVTRTSSYSDSYCMYSGMGTADNLPFGSDPERITWQ 121
 QY 122 YFDROYTASRAVAREVLATGTEPDGVEAFIDCEPLLRFRYPVPEHRSABEOPLEMA 181
 DB 122 YFDSLRYRAAQETARLVLTAGTYDGEDDLTLDCCPVLRLRFPPEVPEHRSABEOPLEMA 181
 QY 182 PHYDLSMTVTLIQOTPCANGFVSLQAEVGAFTDLPPRPDAVLVFCGAIATLVTSQGVKAP 241
 DB 182 PHYDLSMTVTLIQOTPCANGFVSLQAEVGAFTDLPPRPDAVLVFCGAIATLVTSQGVKAP 241
 QY 242 RHHVAAPRRDQIAGSSRTSSVFFLRPMADFTFSPVLARECGFVSLDDEGTATFQOWIGCN 301
 DB 242 NNHVVSPDASMLKSGDRTSSVFFLRPMADFTFSPVLARECGFVSLDDEGTATFQOWIGCN 301
 QY 302 YVNI 305
 DB 302 YVTM 305

RESULT 8
 CEFF_STRCL STANDARD; PRT; 318 AA.
 ID ID CEFF_STRCL
 AC P42220;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE Deacetoxycephalosporin C hydroxylase (EC 1.14.11.-)
 DE (Deacetylcephalosporin C synthetase) (DACS) (Beta-lactam hydroxylase).
 GN Name=ceff;
 OS Streptomyces clavuligerus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1901;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91100311; PubMed=1987130;
 RA Kovacevic S., Miller J.R.;
 RT "Cloning and sequencing of the beta-lactam hydroxylase gene (ceff) from Streptomyces clavuligerus: gene duplication may have led to separate hydroxylase and expandase activities in the actinomycetes.";
 RL J. Bacteriol. 173:396-400 (1991).
 CC -!- FUNCTION: Hydroxylation of deacetoxycephalosporin C in 3'position to form deacetylcephalosporin C.
 CC -!- PATHWAY: Cephalosporin biosynthesis.
 CC -!- SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; M63809; AAA26716.1; -.
 CC PIR; A39204; A39204.
 DR HSSP; P18548; 1RXF.
 DR InterPro; IPR005123; 2OG-Fell_Oxy.
 DR Pfam; PF03171; 2OG-Fell_Oxy; I.
 KW Antibiotic biosynthesis; Iron; Oxidoreductase; Vitamin C.
 SQ SEQUENCE 318 AA; 34584 MW; B17C1CBCE1E67178 CRC64;

Query Match 57.9%; Score 946.5; DB 1; Length 318;
 Best Local Similarity 58.0%; Pred. No. 3.1e-77;
 Matches 181; Conservative 43; Mismatches 80; Indels 3; Gaps 1;

QY 2 DTTVPTFSLAEIQQGLHODEFRRLCKDKLFYLTDCGLTDTLTKSAKDIVIDPFHSGSEA 61
 DB 3 DTPVPIFNLALREGADDEKRECVTGWGVYLTGAGDKDHLRTATDTAMDFANGTEA 62
 QY 62 EKRAVTSVPVPMRGRGFTGLSESTAOITNTGSSYSDSYCMYSGMGTADNLPFGSDPERITWQ 121
 DB 62 EKAAVTTDVPMRGRGYSALAEASTAOVTRTSSYSDSYCMYSGMGTADNLPFGSDPERITWQ 122
 QY 122 YFDROYTASRAVAREVLATGTEPDGVEAFIDCEPLLRFRYPVPEHRSABEOPLEMA 178
 DB 123 YFDSLRYRAAQETARLVLTAGSGYDAEIVGSLDELLDAPVRLRLRFPPEVPEHRSABEOPLEMA 182
 QY 179 RMAPHYDLSMTVTLIQOTPCANGFVSLQAEVGAFTDLPPRPDAVLVFCGAIATLVTSQGV 238
 DB 183 RMAPHYDLSMTVTLIQOTPCANGFVSLQAEVGAFTDLPPRPDAVLVFCGAIATLVTSQGV 242
 QY 239 KAPRHHVAAPRRDQIAGSSRTSSVFFLRPMADFTFSPVLARECGFVSLDDEGTATFQOWIGCN 298
 DB 243 PAPRHHVAPSAGMRGSDRTSSVFFLRPMADFTFSPVLARECGFVSLDDEGTATFQOWIGCN 302
 QY 299 GANYVNI 305
 DB 303 GTNYVTM 309

RESULT 9
 PRELIMINARY; PRT; 319 AA.
 ID ID PRELIMINARY;
 AC Q48740;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Deacetylcephalosporin C synthetase.
 OS Lysobacter lactamgenus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthobacter.
 OX NCBI_TaxID=39596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YK90;
 RX MEDLINE=96287475; PubMed=8703429;
 RA Kimura H., Izawa M., Sumino Y.;
 RT "Molecular analysis of the gene cluster involved in cephalosporin biosynthesis from Lysobacter lactamgenus YK90.";
 RL Appl. Microbiol. Biotechnol. 44:589-596 (1996).
 CC -!- SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase family.
 CC EMBL; X56660; CAA39984.1; -.
 CC PIR; S54100; S54100.
 DR HSSP; P18548; 1RXF.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR GO; GO:0016481; F:oxidoreductase activity; IEA.
 DR GO; GO:0015288; F:porin activity; IEA.
 DR InterPro; IPR005123; 2OG-Fell_Oxy.
 DR Pfam; PF03171; 2OG-Fell_Oxy; I.
 KW Iron; Oxidoreductase; Porin.
 SQ SEQUENCE 319 AA; 35559 MW; 51040CD201B7C272 CRC64;

Query Match 57.7%; Score 944; DB 2; Length 319;
 Best Local Similarity 58.4%; Pred. No. 5.2e-77;
 Matches 178; Conservative 45; Mismatches 82; Indels 0; Gaps 0;

QY 2 DTTVPTFSLAEIQQGLHODEFRRLCKDKLFYLTDCGLTDTLTKSAKDIVIDPFHSGSEA 61
 DB 3 DSGIQLDLEHGVRLDSFRKSLFERGVFYAEDDSITETAKKANDAVMDLFEHSGSAB 62
 QY 62 EKRAVTSVPVPMRGRGFTGLSESTAOITNTGSSYSDSYCMYSGMGTADNLPFGSDPERITWQ 121

Db	1	MTSKVPVPRLLDGLSGKVLTELAAVTTKGIFYLTESGLVDDBHTSARETCVDFPKNGSE	60
Oy	61	AERKAATVSPPVPMRGGTGTGEBSESTAOITNTGSISDYSMCYSMGTADNLPPSCGFELIWT	120
Db	61	EERKAATVTLAABRNARRGFSALFEMESTAVATELGKISDYSTCYSMIGLTFPNRRGFEDVMQ	120
Oy	121	QYPRROYTASRAVARREVLRAATGTBPDG-GVEAFIDCEPLLRFRFPQVPEHRSAAEOPRLR	179
Db	121	DYPRFMTGAAQDVARAVALNVSAGAPLAGESIDDIEFCOPLRLRLRFEPVEPERVAEEBPLR	180
Oy	180	MAPHYDLISMVTLTIOTPCANGFVSLQAENVGAFTDLEPYRPDAVLVFGCAIATLVTSGOVK	239
Db	181	MGPHYDLSITILVHQTMCANCFVSLQCEVOGEFYDLPTLPAAVVFGCAVQTALTATGKVX	240
Oy	240	APRHVAAAPRRDDQIAGSSRTSSVFPLRNADFTTSVFLARECGSDVSLDGTAFQOMIG	299
Db	241	APKRHVSPGHRDQGVGSSSRSTSVEFLRPKPDPFSFNVOOSREMGNVRIPSERITFRMLG	300
Oy	300	GNYVINIRTSKA 311	
Db	301	GNYVMRRDKPA 312	
 RESULT 11 ----- 			
ID	09P4T5	PRELIMINARY;	PRT; 332 AA.
AC	09P4T5;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DR	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DI	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	Deacylcylcephalosporin C synthetase/hydroxylase.		
GN	Name=cefep; Cephalosporium acremonium (Acrcemonium chrysoeum).		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
CC	Hypocrea/cucurbitaceae; Hypocreales; Hypocreaceae; mitosporic Hypocreaceae;		
OX	NCHI_Taxid=5044;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3/2;		
RA	Jekesch K., Nosek J., Kueck U.;		
RL	Submitted (JUL-2000) to the EMBL/GenBank/DDAJ databases.		
CC	-I- SIMILARITY: Belongs to the iron/ascorbate-dependent oxidoreductase family.		
CC	EMBL; AJ040737; CAB96750.1; -.		
DR	HSSP; P18548; IRXG.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0019867; C:outer membrane; IEA.		
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO; GO:0015288; F:portin activity; IEA.		
DR	InterPro; IPR005123; ZOG-Feril_Oxy.		
DR	InterPro; IPR002057; Isopen_N_synth.		
DR	Pfam; PF01171; ZOG-Feril_Oxy; 1.		
DR	PROSITE; PS00185; IPNS_1; UNKNOWN_1.		
KW	PROSITE; PS00186; IPNS_2; UNKNOWN_1.		
KW	Iron; Oxidoreductase; Portin.		
SQ	SEQUENCE 332 AA; 36538 MW; E3DECESE59C28B79 CRC64;		
Query Match 56.5%; Score 92.5; DB 2; Length 332;			
Best Local Similarity 56.7%; Pred. No. 3,2e-75;			
Matches 177; Conservative 43; Mismatches 91; Indels 1; Gaps 1;			
Oy	1	MDTVTPFESIAELIOGHODEFRRCCLRDKGLFYLTDCGLTDTELSAKADIVDFEHQSE	60
Db	1	MTSKVPVPRLLDGLSGKVLTELAAVTTKGIFYLTESGLVDDBHTSARETCVDFPKNGSE	60
Oy	61	AERKAATVSPPVPMRGGTGTGEBSESTAOITNTGSISDYSMCYSMGTADNLPPSCGFELIWT	120
Db	61	EERKAATVTLAABRNARRGFSALFEMESTAVATELGKISDYSTCYSMIGLTFPNRRGFEDVMQ	120
Oy	121	QYPRROYTASRAVARREVLRAATGTBPDG-GVEAFIDCEPLLRFRFPQVPEHRSAAEOPRLR	179
Db	121	DYPRFMTGAAQDVARAVALNVSAGAPLAGESIDDIEFCOPLRLRLRFEPVEPERVAEEBPLR	180
Oy	180	MAPHYDLISMVTLTIOTPCANGFVSLQAENVGAFTDLEPYRPDAVLVFGCAIATLVTSGOVK	239
Db	181	MGPHYDLSITILVHQTMCANCFVSLQCEVOGEFYDLPTLPAAVVFGCAVQTALTATGKVX	240
Oy	240	APRHVAAAPRRDDQIAGSSRTSSVFPLRNADFTTSVFLARECGSDVSLDGTAFQOMIG	299
Db	241	APKRHVSPGHRDQGVGSSSRSTSVEFLRPKPDPFSFNVOOSREMGNVRIPSERITFRMLG	300
Oy	300	GNYVINIRTSKA 311	
Db	301	GNYVMRRDKPA 312	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 19, 2005, 15:06:04 : Search time 24 Seconds
(without alignment)
967.327 Million cell updates/sec

Title: 10719236-1_232-1164

Perfect score: 1636

Sequence: 1 MDTVPTFSIAELQGGHOD.....ATFQDWIGNVNIRTSKA 311

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	99.6	311	3	US-09-413-231-18
2	1210	74.0	313	1	US-08-592-411-15
3	1195.5	73.1	314	1	US-08-592-411-17
4	972.5	59.4	310	5	PCT-US95-04801-6
5	168	10.3	329	3	US-09-413-231-5
6	168	10.3	329	3	US-09-413-231-9
7	165	10.1	335	2	US-08-379-556A-10
8	153	9.4	333	2	US-08-379-556A-8
9	149.5	9.1	327	4	US-09-902-540-15162
10	147	9.0	333	3	US-09-413-231-10
11	143	8.7	347	2	US-08-379-556A-2
12	133.5	8.2	331	3	US-09-413-231-4
13	133.5	8.2	343	3	US-09-454-034-8
14	129	7.9	321	3	US-09-413-231-8
15	122.5	7.5	331	3	US-09-413-231-1
16	122.5	7.5	331	3	US-09-413-231-2
17	122.5	7.5	331	3	US-09-413-231-3
18	122.5	7.5	386	2	US-08-553-367A-2
19	122.5	7.5	386	4	US-09-295-306-2
20	122.5	7.5	386	4	US-09-734-719-2
21	122.5	7.5	386	4	US-09-371-307-78
22	121	7.4	378	2	US-08-553-367A-6
23	121	7.4	378	3	US-09-295-306-6
24	121	7.4	378	4	US-09-734-719-6
25	120.5	7.4	389	4	US-09-248-796A-14285
26	119	7.3	351	4	US-09-489-039A-9415
27	117	7.2	238	4	US-09-270-767-45029

28	116	7.1	344	4	US-09-252-991A-28503	Sequence 28503, A
29	112.5	6.9	377	2	US-08-553-367A-4	Sequence 4, Appli
30	112.5	6.9	377	3	US-09-295-306-4	Sequence 4, Appli
31	112.5	6.9	377	4	US-09-734-719-4	Sequence 20513, A
32	108.5	6.6	390	4	US-09-252-991A-20513	Sequence 7, Appli
33	108	6.6	338	3	US-08-379-556A-4	Sequence 4, Appli
34	107	6.5	241	2	US-08-413-231-6	Sequence 59, Appli
35	105	6.4	338	3	US-09-413-231-5	Sequence 4, Appli
36	104	6.4	321	4	US-09-371-307-59	Sequence 38, Appli
37	103.5	6.3	331	4	US-09-614-912-44	Sequence 40, Appli
38	103.5	6.3	368	4	US-09-614-912-38	Sequence 1631, A
39	103.5	6.3	734	4	US-09-248-796A-16776	Sequence 2, Appli
40	102.5	6.3	296	4	US-09-614-912-40	Sequence 17497, A
41	102.5	6.3	370	4	US-09-614-912-2	Sequence 10, Appli
42	101.5	6.2	403	4	US-09-248-796A-16531	
43	100	6.1	331	4	US-09-719-108-2	
44	96	5.9	453	4	US-09-248-796A-17497	
45	94.5	5.8	112	3	US-09-454-034-10	

ALIGNMENTS

```
RESULT 1
US-09-413-231-18
; Sequence 18, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Wanner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/413,231
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified DAOCS
; NAME/KEY: MUTAGEN
; LOCATION: (181)
; OTHER INFORMATION: Ala181 in native DAOCS modified to Arg
US-09-413-231-18

Query Match          99.6%; Score 1629; DB 3; Length 311;
Best Local Similarity 99.4%; Pred. No. 5e-174;
Matches 309; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDTVPTFSIAELQGGHODFFRCLRDKGFYLTDDGLTDTLKSADLVIDFFEGSE 60
   |||||||
DB 1 MDTVPTFSIAELQGGHODFFRCLRDKGFYLTDDGLTDTLKSADLVIDFFEGSE 60
   |||||||

QY 61 AEKAAVSPVPTMGRGTGLSESTAOITNTGSDYSCMSKMTAANLFPSSGFERIWT 120
   |||||||
DB 61 AEKAAVSPVPTMGRGTGLSESTAOITNTGSDYSCMSKMTAANLFPSSGFERIWT 120
   |||||||

QY 121 QYFPRQYTAASAVAREVLRATGTETPDGVEAFDCEPLRRFRYPQVPEHRSABEQPLRM 180
   |||||||
DB 121 QYFPRQYTAASAVAREVLRATGTETPDGVEAFDCEPLRRFRYPQVPEHRSABEQPLRM 180
   |||||||

QY 181 APHYDLSMVTLLIOOTPCANGFVSLQAEVGAFTDLVPRPAVLYFGAIALTVLGGGVKA 240
   |||||||
DB 181 RPHYDLSMVTLLIOOTPCANGFVSLQAEVGAFTDLVPRPAVLYFGAIALTVLGGGVKA 240
   |||||||

QY 241 PRHHVAAPRRDIOIGSSRTSSVFPLRNADFTFSVPLARECGFVSLDGETATFQDWIGG 300
   |||||||
```

Db 241 PRHHVAAPRRDQIAGSSRTSSVFLLRPVADFTFVPLARECGFDVSLDGETATFQDWIGG 300
QY 301 NYVNIRRTSKA 311
Db 301 NYVNIRRTSKA 311

RESULT 2

US-08-592-411-15

Sequence 15, Application US/08592411

Patent No. 5726032

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Process for the Efficient Production of

TITLE OF INVENTION: 7-ADCA via 2-(Carboxymethylthio)acetyl-7-ADCA and

TITLE OF INVENTION: 3-(Carboxymethylthio)propionyl-7-ADCA

NUMBER OF SEQUENCES: 17

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,411

FILING DATE:

CLASSIFICATION: 435

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 313 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-592-411-15

Query Match 74.0%; Score 1210; DB 1; Length 313;

Best Local Similarity 71.0%; Pred. No. 5,1e-127;

Matches 220; Conservative 40; Mismatches 50; Indels 0; Gaps 0;

QY 2 DTTVPPTSLAELOGGHODEFRRCCLDKGLFYLTDCGL-TDTELKSAKDVIDFEHSGSEA 61
Db 3 DATVPFDLAELREGHQBFRCHREKGVFLYLGKGLAEADHAGREIAVDFPDHGTGEA 62
QY 62 EKRAVTSVPVPMRGRGFTGLSESTAOITNTGSDYSYSCMSMGTDNLFPSSGDFPERIWTQ 121
Db 63 EKXAVMTPIPTIRRGYAGLSESTAOITNTGKYTDYSMSMGTDNLFPSEAEKXAMED 122
QY 122 YFDRQYTSRAVAAREVLATGTEPDGGEAFLLDCEPLLRFRYPQVPEHRSAGEQPLRMA 181
Db 123 YFARMYRASQDVARQVLTSGAEPEVGMDFLLDCEPLLRFRYPPEVEDRVAEEQPLRMA 182
QY 182 PHYLSMWTLLIQOTPCANGFVSLQAEVGAFTDL.PYRPDAVLVFCGAIATLVTTGGVAKAP 241
Db 183 PHYLSITVTLHQTPCANGFVSLQAEVGVSDVDIPAGQAVLVFCGAVALTVADGAIKAP 242
QY 242 RHHAAPRRDQIAGSSRTSSVFLLRPVADFTFVPLARECGFDVSLDGETATFQDWIGGN 301
Db 243 KHHVAAPGADKRVGSSRTSSVFLLRPVADFTFVPLARECGFDVSLPATAITFDMDWIGN 302
QY 302 YVNIRRTSKA 311
Db 303 YINIRKTAATA 312

RESULT 3

US-08-592-411-17

Sequence 17, Application US/08592411

Patent No. 5726032

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Process for the Efficient Production of

TITLE OF INVENTION: 7-ADCA via 2-(Carboxymethylthio)acetyl-7-ADCA and

TITLE OF INVENTION: 3-(Carboxymethylthio)propionyl-7-ADCA

NUMBER OF SEQUENCES: 17

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,411

FILING DATE:

CLASSIFICATION: 435

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 314 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-592-411-17

Query Match 73.1%; Score 1195.5; DB 1; Length 314;

Best Local Similarity 70.4%; Pred. No. 2,2e-125;

Matches 219; Conservative 40; Mismatches 51; Indels 1; Gaps 1;

QY 2 DTTVPPTSLAELOGGHODEFRRCCLDKGLFYLTDCGL-TDTELKSAKDVIDFEHSGSEA 60
Db 3 DATVPFDLAELREGHQBFRCHREKGVFLYLGKGLAEADHAGREIAVDFPDHGTGEA 62
QY 61 AEKRAVTSVPVPMRGRGFTGLSESTAOITNTGSDYSYSCMSMGTDNLFPSSGDFPERIWT 120
Db 63 AEKXAVMTPIPTIRRGYAGLSESTAOITNTGKYTDYSMSMGTDNLFPSEAEKXAMED 122
QY 121 QYFDRQYTSRAVAAREVLATGTEPDGGEAFLLDCEPLLRFRYPQVPEHRSAGEQPLRM 180
Db 123 DYFARMYRASQDVARQVLTSGAEPEVGMDFLLDCEPLLRFRYPPEVEDRVAEEQPLRM 182
QY 181 APHYLSMWTLLIQOTPCANGFVSLQAEVGAFTDL.PYRPDAVLVFCGAIATLVTTGGVAKAP 240
Db 183 APHYLSITVTLHQTPCANGFVSLQAEVGVSDVDIPAGQAVLVFCGAVALTVADGAIKAP 242
QY 241 PRHHVAAPRRDQIAGSSRTSSVFLLRPVADFTFVPLARECGFDVSLDGETATFQDWIGG 300
Db 243 KHHVAAPGADKRVGSSRTSSVFLLRPVADFTFVPLARECGFDVSLPATAITFDMDWIGG 302
QY 301 NYVNIRRTSKA 311
Db 303 NYINIRKTAATA 313

RESULT 4

PCT-US95-04801-6

Sequence 6, Application PC/TUS9504801

GENERAL INFORMATION:

APPLICANT: Martin, Juan F.

APPLICANT: Coque, Juan R.

APPLICANT: Enguita, Francisco J.

APPLICANT: Fuente, Juan L.

APPLICANT: Lirias, Francisco J.

APPLICANT: Lirias, Paloma

TITLE OF INVENTION: DNA ENCODING CEPHAYNCIN BIOSYNTHESIS

TITLE OF INVENTION: LATE GENES

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: John W. Wallen III

STREET: P.O. Box 2000

CITY: Rahway

STATE: New Jersey

COUNTRY: USA

ZIP: 07065

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04801

```
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Wallen III, John W.
/ REGISTRATION NUMBER: 35,403
/ REFERENCE/DOCKET NUMBER: 19179
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 594-3905
/ TELEFAX: (908) 594-4720
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 310 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ PCT-US95-04801-6

Query Match          59.4%; Score 972.5; DB 5; Length 310;
Best Local Similarity 61.5%; Pred. No. 2,2e-100;
Matches 187; Conservative 37; Mismatches 79; Indels 1; Gaps 1;

QY 2 DTTVPFSLAELOQGLHODEFRRCILDKGLFYLTDCGLTDTLKSAXDIVIDPFPHGSA 61
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 3 DKTVPFVSMALRDSRQDEFRMAR-RGVFYLTGATERDHRVATDTAMDFFAQGTAE 61
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 62 EKRATSPVPTMRGRFTGLESESTAOITNTGSDYSYSCMYSGTADNLPFGSDFERTIQ 121
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 62 EKQAVTTRKYPTMRKGSALAEASTAOITNTGYSMSYSGTADNLPFGSDFERTIQ 121
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 122 YEDROYTASRAVEYLRTATGTEPDGVAFDCEPILRFYFPOVPEHRSABEOLPMA 181
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 122 YEDSLYRAAQERANLVLTAAGTYDGEDDLTLDCCPVLRLRFPVPEHRAAEYEPKMA 181
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 182 PHYDLSMTLLIQOTPCANGFVSLQAEVGAFTDLPYRPDAVLVFCGAIATVITGQVAP 241
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 182 PHYDLSITFIHQTPCANGFVSLQAEVDEMYSLPHVEDAVVVLGAIAPLVITGQAVAP 241
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 242 RHVVAAPRRDQIAGSSRTSVFLLRPNADFTSVPLARCGSDVLDGTAFTQWIGN 301
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 242 NHVAVSPDASMLKSGDRTSVFLRPSTDPFTSVDPARKYGLVDLMEKATFGWIGTN 301
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 302 YVNI 305
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 302 YVTM 305
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 5
US-09-413-231-5
/ Sequence 5, Application US/09413231
/ Patent No. 6284483
/ GENERAL INFORMATION:
/ APPLICANT: Dilley, David R
/ APPLICANT: Kadyrzhanova, Dina K
/ APPLICANT: Wang, Zhenyong
/ APPLICANT: Warner, Toni M
/ TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
/ TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
/ FILE REFERENCE: MSU41-453
/ CURRENT APPLICATION NUMBER: US/09/413,231
/ CURRENT FILING DATE: 1999-10-06
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
/ NAME/KEY: MUTAGEN
/ LOCATION: (210)
/ OTHER INFORMATION: Glu210 in native IPNS modified to Arg
```

```
US-09-413-231-5
Query Match          10.3%; Score 168; DB 3; Length 329;
Best Local Similarity 23.8%; Pred. No. 4.2e-10;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAE-----OQGLHODEFRRCILDKGLFYLTDCGLTDTLKSAXDIVIDPFPHG 58
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 10 VPTIDISLPLFTDAAKRAVEIHGACRSGGFYATNHGV---DVQQLDVNVEFHGM 66
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 59 SEAEK-----RAVTSVPVPTMRGRF---TGLESESTAOITNTGSDYSYSCMYSGT---A 106
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 67 TDQEKHLAIHAAYPNPDHVNANGYKAVPKGKAVSEFCYLPNDPGEDEHPM-IAAGTPEH 125
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 107 DNLFPSSG---FERITQYDROYTASRAVEYLRTATGTEPDGVAA-----FLDCE 156
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 126 VNLMPDEERHPRFPCEGYRQMLKSTVIMRGIALALG-RPHFPDAALAEODSLSSV 184
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 157 PLRFRYPPOVPEHRSABE-OLPMAPHYDLSMTLLIQOTPCANGFVSLQAEVGAFTDL 215
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 185 SLIRPYLAEYFPYKTPGDGLSFRDHDVSMITVLFTQYQV---LOVETVDGMRDI 240
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 216 PYRPDAVLVFCGAIATVITGQVAPRHHVVAAPRRDQIAGSSRTSVFLLRPNADFTSV 275
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 241 FTSENDPLVNGCYMAHVTNDYPPAPNHRV-----KFNVERLSLPFLNGCHBAVIR- 293
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 216 PLARECGPDVLDGETATFQDWI 298
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 294 PFVPE-GASEVRNEALSYGDL 315
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 6
US-09-413-231-9
/ Sequence 9, Application US/09413231
/ Patent No. 6284483
/ GENERAL INFORMATION:
/ APPLICANT: Dilley, David R
/ APPLICANT: Kadyrzhanova, Dina K
/ APPLICANT: Wang, Zhenyong
/ APPLICANT: Warner, Toni M
/ TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
/ TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
/ FILE REFERENCE: MSU41-453
/ CURRENT APPLICATION NUMBER: US/09/413,231
/ CURRENT FILING DATE: 1999-10-06
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 9
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
/ NAME/KEY: MUTAGEN
/ LOCATION: (210)
/ OTHER INFORMATION: Glu210 in native IPNS modified to Arg

US-09-413-231-9
Query Match          10.3%; Score 168; DB 3; Length 329;
Best Local Similarity 23.8%; Pred. No. 4.2e-10;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

QY 5 VPTFSLAE-----OQGLHODEFRRCILDKGLFYLTDCGLTDTLKSAXDIVIDPFPHG 58
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 10 VPTIDISLPLFTDAAKRAVEIHGACRSGGFYATNHGV---DVQQLDVNVEFHGM 66
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 59 SEAEK-----RAVTSVPVPTMRGRF---TGLESESTAOITNTGSDYSYSCMYSGT---A 106
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 67 TDQEKHLAIHAAYPNPDHVNANGYKAVPKGKAVSEFCYLPNDPGEDEHPM-IAAGTPEH 125
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 107 DNLFPSSG---FERITQYDROYTASRAVEYLRTATGTEPDGVAA-----FLDCE 156
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Db 126 VNLMPDEHRHRRPFCBGYYRQMLKLTVMRLGALALG-RPHEFDALALBODSLSSV 184
Qy 157 PLRFRFRPPQVPEHRSAEE-OLPLMABHYDLISMTLLIQOTPCANGFVSLQAEVCGATDL 215
Db 185 SLIRPYLAEERYPKVTGPDGLSFRDLHDVSMITVLFQTOVQN----LQVETDVGWRDI 240
Qy 216 PYRDAVALVFCGALATLVTTGGQVAKPRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSV 275
Db 241 PTSNDPLVNCGTMAHYTNDYPPAPNHRV-----KFAVNERLSLFFPLANGHEAVIE- 293
Qy 276 PLARECGFVSLDEGTATFQDWM 298
Db 294 PFVPE-GASEVRNEALSVDYL 315

RESULT 7
US-08-379-556A-10
; Sequence 10, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:
; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-379-556A-10

Query Match 10.1%; Score 165; DB 2; Length 335;
Best Local Similarity 22.8%; Pred. No. 9.3e-10;
Matches 72; Conservative 55; Mismatches 135; Indels 54; Gaps 14;

Qy 5 VPTFSIAELQOGLHODEFRRCCLRDKGLFYLTDCGLTDTLTKSAKDVIYDPFHSSEAKR 64
Db 43 VPVIDLSQADNESLVALISKASKDWGIFQVNNHGISPLISKLOVKGKFFELFQF-EKE 101
Qy 65 AVTSPVPTMRGRGTGLSESTQAQITNTGSSYSDYSCWYSGMTADNLF-----PSGDFERIW 119
Db 102 VIAPR-----DGYGVGVE-----GYGTKLQKEVQGGKGVNDLPHLWMPSPSTINQFW 148
Qy 120 TQ-----YFD-RQYTAAR-AVAREVL-----RATGTEPD-----GCVBAFLDCPPLRFR 162
Db 149 PKNPSPYRDTNEEYTSQSLIGVANKLLGLSLKGLGEDEVKQALGGBDLIY---MDKIN 204

Qy 163 YRPQVPEHRSABEQPLRMAPHYDLISMTVLLIQOTPCANGFVSLQAEVCGATDLPYRDAV 222
Db 205 YTPPCP-----CEPLALGVAPHDTMSSITLL-----VNEVQGLQVFPDGGQYDVATYPMNL 256
Qy 223 LVECGAIATLVTTGGQVAKPRHHVAAPRRDQIAGSSRTSSVFPLRPNADFTFSVPLARECG 282
Db 257 IHHGQIEILSNGKYSVYHRSTYVKE-----KTRMSWPAFLPEPPEVGG-PRLKLVN 310
Qy 283 FQVSLDEGTATFQDWM 298
Db 311 KDDPPRYKTKYKXDYV 326

RESULT 8
US-08-379-556A-8
; Sequence 8, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:
; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9592
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)742-4343
; TELEFAX: (516)742-4366
; TELEEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-379-556A-8

Query Match 9.4%; Score 153; DB 2; Length 333;
Best Local Similarity 21.6%; Pred. No. 2e-08;
Matches 71; Conservative 57; Mismatches 126; Indels 74; Gaps 15;

Qy 1 MDTT-VPTFSIAELQOGLHODEFRRCCL-----RDKGLFYLTDCGLTDTLTKSAKDVIYDP 54
Db 39 LDTVLVEPALDLS-----LEEDDVVKLVLSASKEMGLFQVTNNGIPTVEYKELQVKGKMF 93
Qy 55 FEHGSABEKRAVTSPPVPTMRGRGTGLSESTQAQITNTGSSYSDYSCWYSGMTADNLFPSGD 114
Db 94 FRAPAR-EKERTIKPKGVEGVGTMLQKEIQGR-----KGVVDHL----- 132
Qy 115 FERITWT-----QYF-----DRQYASRAVAREVL-----RATGTEPDG-----GV 149
Db 133 FHKVMPSPVNTYNNWTKTSYRANAEYTKYLRIVADKLFKCKSKGLGLEDEVKSKCGN 192
Qy 150 EAFIDCEPPLRFRFYFPQVPEHRSABEQPLRMAPHYDLISMTVLLIQOTPCANGFVSLQAEV 209

```
Db      193 E----DIVYLKINYYPCPR-----PDIALGVAHTDLSVITL-----VENDVAGLQVSRD 241
Qy      210 GAFDLPYRPDAVAVCCGAIATLVGGGVKAPRHVAAPRRQIAGSSSTSSVFLRPA 269
Db      242 GEMVDKTYIENALIIHVGGQMEIMSGEYKVALHSTVKE-----RTIRISVPEVLEPPS 296
Qy      270 DFTFSVPLARECGFVSLDGETATFQDM 297
Db      297 DPAVG-PIPKLISDEKPAKTKIVSEY 323

RESULT 9
US-09-902-540-15162
; Sequence 15162, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15162
; LENGTH: 327
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-15162

Query Match      9.1%; Score 149.5; DB 4; Length 327;
Best Local Similarity 21.6%; Pred. No. 4.9e-08;
Matches 77; Conservative 39; Mismatches 118; Indels 123; Gaps 15;

Qy      21 EERRCLRDKGLFYLTDCGLDTTEL-----KSAADIYI-----DFPEH 57
Db      19 EIEQCRDSGFFYVGHGVSGLVLRLERESHRFPALPVAKAELAMSGVAMRGWPL 78
Qy      58 GSEAEKRAVTSVPVPMRKG-FTGLESESTAOITNTG-----SYSDY- 97
Db      79 GGE-----LTSGRPRKRGSLYGTLESGHPRVXKGMPLHGANLMPAEVPELRAVLDYV 133
Qy      98 -----SVCYSMGTADNLPFSGDPERIMTQYPDROYTASRAVAREVLRATGEP 145
Db      134 AACTRAAHALMEGMALSLG-----LDADYFRRHHTADPTV----- 168
Qy      146 DGGVAFPLDCEPLRFRRYPPQVBEHRSABEQPLRMAHYDLSMTLLIOQTPCANGFVSLQ 205
Db      169 -----LRFIFHPAEPOH---EVSWSGGEHTDGLTLTLQDD--NG--GLQ 209
Qy      206 AEVGGAFTDLPYRPDAVAVFCGAIATLVGGGVKAPRHVAAPRRDQIAGSSSTSSVFL 265
Db      210 VTPRGWVEVPLPGTLVGNIGMDLDMTCGMYRSTPHRV-----KNVSGKDRLSFPLFF 264
Qy      266 RPNADFTTSV-PLAEECGDV-----SLDGETATFQDMIGVYVIRRTSK 310
Db      265 DP--DFAAEVHPLPRGAGADVDDDRARRWDGASVHAFQGTGYDYLGVFKVLSRSR 319

RESULT 10
US-09-413-231-10
; Sequence 10, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dille, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
```

```
; FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/413,231
; CURRENT FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 333
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; NAME/KEY: MUTAGEN
; LOCATION: (214)
; OTHER INFORMATION: Glu214 in native IPNS modified to Arg
US-09-413-231-10

Query Match      9.0%; Score 147; DB 3; Length 333;
Best Local Similarity 23.2%; Pred. No. 9.6e-08;
Matches 70; Conservative 37; Mismatches 131; Indels 64; Gaps 13;

Qy      5 VPTSLAELOQGLHOD-----EERRCLRDKGLFYLTDCGLDTTELKSAADIYIDPEH 57
Db      10 VPTIDISPL-FGTPDPAKAHVAQINEACRSGCFEYASHHG1---DVRLODVNVEFRT 65
Qy      58 GSEAEKRAVTSVPVPMRKGFTGLESESTAOITNTGSDYSDYSCY---SMGTADNL 109
Db      66 MTQDEKIDLAIHAYNENNSHVRNGY-----YMARPGKTVESWCYLNPSGEDHPM 116
Qy      110 FPGS-----DPERIMTQYPDROYTASRA--VAREVLRATG-----EPDG 147
Db      117 IKAGTPMHEVNVWPDDEHRHDPFRSGQYREVRRLKVLRLRFPALALGPEEFENEV 176
Qy      148 GVEAFDCEPL-LRFYFPQVPE--HRSABEQPLRMAHYDLSMTLLIOQTPCANGFVS 203
Db      177 TEEDTLSCRSIMIRYPLDYPPEAIIKGTGDTRLSPRDHLDVSMITVLFGTEVON---- 232
Qy      204 LOAEVGAFTDLPYRPDAVAVFCGAIATLVGGGVKAPRHVAAPRRDQIAGSSSTSSV 263
Db      233 LVETVDGMOSSLPTSGENFLNCGTLYGYLTNDYFPAPNHRV-----KYNAERLSLP 286
Qy      264 FL 265
Db      287 FL 288

RESULT 11
US-08-379-556A-2
; Sequence 2, Application US/08379556A
; Patent No. 5859329
; GENERAL INFORMATION:
; APPLICANT: HOLTON, TIMOTHY A.
; APPLICANT: KEAM, LISA A.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
; TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,556A
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
```

```

NAME: DIGITLIO, FRANK S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9592
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)742-4343
TELEFAX: (516)742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-379-556A-2

```

[illegible]

```

RESULT 12
US-09-413-231-4
; Sequence 4, Application US/09413231
; Patent No. 6284483
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MS041-453
; CURRENT APPLICATION NUMBER: US/09/413,231
; CURRENT FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Penicillium chrysogenum
; NAME/KEY: MUTAGEN
; LOCATION: (212)
; OTHER INFORMATION: Glu212 in native IPNS modified to Arg
; US-09-413-231-4

```

```

Query Match      8.2%; Score 133.5; DB 3; Length 331;
Best Local Similarity 23.8%; Pred. No. 3.1e-06;
Matches 76; Conservative 40; Mismatches 131; Indels 73; Gaps 14.
0y 27 RDKGIFLYITDCG-----LTTDELSKANDIVDF--EHGSE----- 60

```

Db 37 RDGGFFVAHVHGVDVGRKLSNKTREHFSTIDEE---KMDLAIIRYAKHEHODQIACGYTIS 93

QY 61 -AEKRAVTSVPMTMRGFTGLESBSTAQITNTGYSIDYSMCYSMGTA--NLPS-----G 113

Db 94 IPEKKAVES-----FCYL-----NPNFKDHPHLIQSKTPHEHVNVMDDEKKHP 136

QY 114 DFERITQTFDRQYITSRVAAREVLRATGTEPDGGEAF-----LDCEPLIRFFYFQVR 168

Db 137 GFEEFAEQYYMDVFGGSLALGRYALAGKEEDPFSRHFKEDALLSSVILIRPYLNP 196

QY 169 EH--RSAER-QPLRMAPHYDLSMTLLIOOTPCANGFVSLQAEVGCAETDLPYREDAVLVE 225

Db 197 PAIKIAEDQTKLSPFMHEDVSLITVLVQSDVAN---LQYEMQGVIQDIEADNATIVN 252

QY 226 CGAIIATLVTCGGQVKAQRHHVVAAPRRDQIAGSSRTSSVFFLRPNADPTFSPVPLARECGFDV 285

Db 253 CGSYMAHITNNYYPAPIRHFV-----KWNNEERQSLPFV--NLGFNDVQVPMWPSKEDG 304

QY 286 SLDGETATPDQWIGGNYVNI 305

Db 305 KTDQRPISYGDYDQNGVLVL 324

```

RESULT 13
US-09-454-034-8
; Sequence 8, Application US/09454034
; Patent No. 6380464
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Fader, Gary M.
; TITLE OF INVENTION: Plant Flavonol Synthase Homologs
; FILE REFERENCE: B11298 US NA
; CURRENT APPLICATION NUMBER: US/09/454,034
; EARLIER FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,875
; NUMBER OF SEQ. ID NOS.: 10
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Glycine max
US-09-454-034-8

```

Query Match	8.2%	Score 133.5	DB 3	Length 343
Best Local Similarity	22.7%	Pred. No. 3.3e-06		
Matches	65	Conservative 46	Mismatches 126	Indels 49
Gaps				133
QY	5	VPTFSLAELOOGLHODEFRCLRDKGLFYLTDCGLTDTLKLASARDIIVDFPFHSGSEAKR	64	
DB	51	VPLTIDSDPDDEGKGVHIELEASRDGMQFIIVHDIIPSDVIKLOSVMGFEEFLPOE-EKE	109	
QY	65	AVTSPVPTRRRGFTGLSESESTAQIINTTSYSYDSMCSYMGADNLNF-----PSCGFERIW	119	
DB	110	LIAP-----AGSDS-----IEGYFKLOKEVNGKKGVNDHLFHWPPSSINISFW	156	
QY	120	TQ-----YFDROYTAS-PAVAREVLA-----TGTE-----PDGVEAFLDCEPLIRRYF	164	
DB	157	PONPPSYREVNVEYSCHLRGVVDKLFKMSVSGLEENELKEGNED--DMWYLLKLNYY	214	
QY	165	POVPEHRSAEQELRYMAPHYDLSMTYLLIQTPFCAGFSLQAEVGAFTDLPYRPDAVLV	224	
DB	215	PPCP-----CEDVLVSGVPHITDMSYLTIL-----VPNEVGGLOACRDMHYDVKYVPNALVI	266	
QY	225	FCGALATLVGGGVKAPRHHVAAPRRDOIAGSSRTSSVFLIRPNAD	270	
DB	267	HIDQWELISLNGKYAKFERRTV-NKDE-----TRMSWVFLEPKE	307	

RESULT 14
US-09-413-231-8
; Sequence 8, Application US/09413231

```
/ Patent No. 6284483
/ GENERAL INFORMATION:
/ APPLICANT: Dilley, David R
/ APPLICANT: Kadyrtzanova, Dina K
/ APPLICANT: Wang, Zhenyong
/ APPLICANT: Warner, Toni M
/ TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
/ TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
/ FILE REFERENCE: MSU41-453
/ CURRENT APPLICATION NUMBER: US/09/413,231
/ CURRENT FILING DATE: 1999-10-06
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 8
/ LENGTH: 321
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
/ OTHER INFORMATION: from Streptomyces cattleya
/ NAME/KEY: MUTAGEN
/ LOCATION: (211)
/ OTHER INFORMATION: Glu211 in native IPNS modified to Arg
US-09-413-231-8

Query Match          7.9%; Score 129; DB 3; Length 321;
Best Local Similarity 22.8%; Pred. No. 9.5e-06;
Matches 72; Conservative 32; Mismatches 150; Indels 62; Gaps 12;

QY 5 VPTFSIABIQQGLHODEFR-----RCLRDKGLFYLTDCGLTDELKSAKDVIDFEH 57
DB 10 VPTIDSPOLFGTDDPPRRTSRGRSTRPARSGGFYASHHGIDVRLQWSN----- 61
QY 58 GSEAEKRAVTSVPVTRRCFTGLSESTAQ---ITNGSYSDYSMCY--SMGTADNLF 110
DB 62 ---ESTWTQDRSTWRSTRYNNENSHVNGYVARPERETVESGYNLPSFGEDHPMM 117
QY 111 PSG-----DPERIMTQYDRQYTSRAV---AREVLRAITGNEPRGCV-- 149
DB 118 KAGTGMHEVNVWPDERRHPDFSGEQY-HREVSASRCCCGASRRRQAGSSSNEVTE 176
QY 150 EAFLDCEPLLRFRYPQVPE---HRSABEQPLRMAPHYDLSMTLLIOQTPCANGFVSLQA 206
DB 177 EDTLSAVSMIRYPYLDYPEALIKTGPDGTRLSFRDHLVSMITVLSKTEYQN----LQY 232
QY 207 EVGGAFTDLPYRPDAVLVFCGAIATLVGQYKAPRHVVAAPRDQIAGSSRTSSVFELR 266
DB 233 ETVVDGQSLPTSGENFLINCGTYLGTINDYFPAPNHRV-----KYVNAERLSLPFLH 286
QY 267 PNADFTFSVPLARECG 282
DB 287 AGQNSVMK-PFTRRTG 301

RESULT 15
US-09-413-231-1
/ Sequence 1, Application US/09413231
/ Patent No. 6284483
/ GENERAL INFORMATION:
/ APPLICANT: Dilley, David R
/ APPLICANT: Kadyrtzanova, Dina K
/ APPLICANT: Wang, Zhenyong
/ APPLICANT: Warner, Toni M
/ TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
/ TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
/ FILE REFERENCE: MSU41-453
/ CURRENT APPLICATION NUMBER: US/09/413,231
/ CURRENT FILING DATE: 1999-10-06
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 331
/ TYPE: PRT
```

```
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
/ OTHER INFORMATION: from Emericella nidulans strain D1A1
/ NAME/KEY: MUTAGEN
/ LOCATION: (212)
/ OTHER INFORMATION: Glu212 in native IPNS modified to Arg
US-09-413-231-1

Query Match          7.5%; Score 122.5; DB 3; Length 331;
Best Local Similarity 20.9%; Pred. No. 5.4e-05;
Matches 68; Conservative 45; Mismatches 129; Indels 83; Gaps 13;

QY 27 RDKGLFYLTDCGLTDELKSAK-----DIVIDPF--EHGSEAEKRAVTSVPV 71
DB 37 RDTGFFVAVNHGINVQRLSQKTKEFHMSITPEEKMDLAIRAYNKEHQDQ----- 85
QY 72 TMRAGFTGLSESTAQITNTGSYSDYSMCY--SMGTAD-----NLPS----- 112
DB 86 -VRAGY-----YLSIPGKKAVESFCYLNPNFTPDHRIQAKTPTHEVNVWPDERTKH 135
QY 113 GDERIMTQYDRQYTSRAVAREVLRAITGE-----PDGVEAFLDCEPLLRPY 163
DB 136 PGFQDFAEQYWDVDFGLSSALLKGYALALGKENFFARHFRPDPTLASV-----LIRYPY 191
QY 164 FPQVPE---HRSABEQPLRMAPHYDLSMTLLIOQTPCANGFVSLQAEVGAFTDLPYRPD 220
DB 192 LDYYPEAIIKTAADGTLSFRMHBDVSLITVLYQSNVQN---LQVETAGYQDI EADDT 247
QY 221 AVLVECGAIATLVGQYKAPRHVVAAPRDQIAGSSRTSSVFELRPNADFTFSVPLARE 280
DB 248 GYLINGCSYMAHLTNNTYKAPIHRV-----KWNNAERQSLPFVNLGYDSVIDPEPDRE 301
QY 281 CGFDVSLDGETATFQDWIGNYVNI 305
DB 302 P--NGKSDRBLSYGDYLLQNGLVS L 324
```

Search completed: September 19, 2005, 15:19:20
Job time : 25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 19, 2005, 15:16:46; Search time 85 Seconds
(without alignments)
1481.559 Million cell updates/sec

Title: 10719236-1_232-1164

Perfect score: 1636

Sequence: 1 MDVTVPFTSLAEIQGLQHD.....ATFDWIGMGNVYINRTSKA 311

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	1629	99.6	311	US-09-924-841-18	Sequence 18, Appl
2	173.5	10.6	345	US-10-732-923-9965	Sequence 9665, Ap
3	171	10.5	342	US-10-732-923-9978	Sequence 9978, Ap
4	169	10.3	366	US-10-369-493-12239	Sequence 12239, A
5	169	10.3	366	US-10-732-923-10056	Sequence 10056, A
6	168	10.3	329	US-09-924-841-5	Sequence 5, Appl
7	168	10.3	329	US-09-924-841-9	Sequence 9, Appl
8	168	10.3	412	US-10-732-923-9964	Sequence 9964, Ap
9	156	9.5	375	US-10-425-115-22033	Sequence 22033, A
10	156	9.5	375	US-10-425-114-58493	Sequence 58493, A
11	155	9.5	352	US-10-425-114-63175	Sequence 63175, A

12	154.5	9.4	366	US-10-437-963-146275	Sequence 146275,
13	154	9.4	342	US-10-425-115-22032	Sequence 22032,
14	154	9.4	371	US-10-425-114-62424	Sequence 62424, A
15	154	9.4	372	US-10-425-114-53454	Sequence 53454, A
16	154	9.4	372	US-10-425-114-72059	Sequence 72059, A
17	153	9.4	350	US-10-437-963-136698	Sequence 136698,
18	148.5	9.1	327	US-10-369-493-14043	Sequence 14043, A
19	148.5	9.1	344	US-10-425-115-316391	Sequence 316391,
20	148.5	9.1	356	US-10-424-599-258352	Sequence 258352,
21	147	9.0	333	US-09-924-841-10	Sequence 10, Appl
22	146.5	9.0	348	US-10-437-963-143438	Sequence 143438,
23	145	8.9	362	US-10-437-963-181202	Sequence 181202,
24	144.5	8.8	382	US-10-437-963-148589	Sequence 148589,
25	144.5	8.8	389	US-10-481-381-4	Sequence 4, Appl
26	143.5	8.8	342	US-10-757-701-44976	Sequence 44976, A
27	143	8.7	359	US-10-136-444-8	Sequence 8, Appl
28	140.5	8.6	428	US-10-437-963-176920	Sequence 176920,
29	139.5	8.5	380	US-10-431-273-88	Sequence 88, Appl
30	139.5	8.5	386	US-10-425-115-347325	Sequence 347325,
31	136	8.3	611	US-10-437-963-104090	Sequence 104090,
32	135.5	8.3	259	US-10-369-493-17939	Sequence 17939, A
33	134.5	8.2	405	US-10-739-930-7737	Sequence 7737, Ap
34	134	8.2	341	US-10-437-963-140887	Sequence 140887,
35	133.5	8.2	331	US-09-924-841-4	Sequence 4, Appl
36	133.5	8.2	334	US-10-424-599-190993	Sequence 190993,
37	133.5	8.1	350	US-10-425-115-254679	Sequence 254679,
38	131.5	8.0	339	US-10-369-493-7649	Sequence 7649, Ap
39	131	8.0	334	US-10-732-923-10146	Sequence 10146, A
40	131	8.0	356	US-10-425-114-64811	Sequence 64811, A
41	131	8.0	356	US-10-739-930-7228	Sequence 7228, Ap
42	130.5	8.0	353	US-10-425-114-61294	Sequence 61294, A
43	130.5	8.0	356	US-10-424-599-259801	Sequence 259801,
44	130.5	8.0	366	US-10-424-599-211485	Sequence 211485,
45	130.5	8.0	382	US-10-425-114-66617	Sequence 66617, A

ALIGNMENTS

RESULT 1

US-09-924-841-18

Sequence 18, Application US/09924841

Patent No. US20020127633A1

GENERAL INFORMATION:

APPLICANT: Dille, David R

APPLICANT: Kadyrzhanova, Dina K

APPLICANT: Wang, Zhenyong

APPLICANT: Warner, Toni M

TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and

TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate

FILE REFERENCE: MS041-453

CURRENT APPLICATION NUMBER: US/09/924, 841

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US/09/413, 231

PRIOR FILING DATE: 1999-10-06

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 18

LENGTH: 311

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: modified DAOCS

OTHER INFORMATION: from Streptomyces clavuligerus

NAME/KEY: MUTAGEN

LOCATION: (181)

OTHER INFORMATION: Ala181 in native DAOCS modified to Arg

US-09-924-841-18

Query Match 99.6%; Score 1629; DB 9; Length 311;
Best Local Similarity 99.4%; Pred. No. 1.9e-166;
Matches 309; Conservative 1; Mismatches 1; Gaps 0;

```

RESULT 3
US-10-732-923-9978
; Sequence 9978, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796) C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 2419
; SEQ ID NO 9978
; LENGTH: 342
; TYPE: prt
; ORGANISM: Magnetospirillum magnetotacticum
US-10-732-923-9978

Query Match      10.5%; Score 171; DB 17; Length 342;
Best Local Similarity 26.3%; Pred. No. 3e-09;
Matches 87; Conservative 38; Mismatches 154; Indels 52; Gaps 14

```

[illegible]

```

RESULT 4
US-10-369-493-12239
/ Sequence 12239, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ. ID NOS: 47374
/ SEQ. ID NO. 12239
/ LENGTH: 366
/ TYPE: PRT
/ ORGANISM: Mesorhizobium loti
US-10-369-493-12239

```

Best Local Similarity 24.7%; Pred. No. 5.4e-09;
Matches 85; Conservative 41; Mismatches 158; Indels 60; Gaps 14;

```
QY 1 MDTTPTFSLAIQGLHQ-----DEFRCLDKGLFYLTDCGLDTBELKSAKDVIDPF 55
DB 28 MPRIVFVLDSRLLEGASERRTFPLDLRSASRDIGFYLAGHGISWAEISEVLTASRQPF 87
QY 56 HNGSEAEKRAVTSVPVTRMGFTGLESSESTAQITNTGSSDYSCMCSMGTAANLFPSSGPF 115
DB 88 AL-PEADKLAIEWVSSQFRGTYRAGGELTK-----GREDMREQLDIVEROALIAQGGG 140
QY 116 ERIWTO-YFDROYTAS-----RAVAEVLNANGTEBPDGVEAFIDCEPIL 159
DB 141 TPAMTRLQGPQNPALPDLKALLAWOSKVTAVALRLKAFQSLDQPEDAF---DPI- 196
QY 160 RRRYFQVDEHR-SAEEOPLR-----MAPHYDISMTLLIQPTPCANGFVSLQAEVGG 210
DB 197 ---YSSBNHMKIVRYRGRDITGGDQGVGAHKDGGFLLTLQ---DNNKGLQVDYDG 248
QY 211 AFTDLPRPDVAVFCGATATLVGGQVAPRHHVAAPRRDQIAGSSRTSSVFLRPND 270
DB 249 SWVDVPIGTIVNIGELLEASNGYLAIVHVVTP---AGVERISVFFFSARLD 304
QY 271 FT-----FSVPLARECGFVSLDGETATFQDWIGGNYVNIIRTS 309
DB 305 ATIPILGSEELAQAARGPAS-DPDNPLFRD-VGTNVLKSRLRS 346
```

```
RESULT 5
US-10-732-923-10056
; Sequence 10056, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgeton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10056
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mesozhizodium loti
US-10-732-923-10056
```

Query Match 10.3%; Score 169; DB 17; Length 366;
Best Local Similarity 24.7%; Pred. No. 5.4e-09;
Matches 85; Conservative 41; Mismatches 158; Indels 60; Gaps 14;

```
QY 1 MDTTPTFSLAIQGLHQ-----DEFRCLDKGLFYLTDCGLDTBELKSAKDVIDPF 55
DB 28 MPRIVFVLDSRLLEGASERRTFPLDLRSASRDIGFYLAGHGISWAEISEVLTASRQPF 87
QY 56 HNGSEAEKRAVTSVPVTRMGFTGLESSESTAQITNTGSSDYSCMCSMGTAANLFPSSGPF 115
DB 88 AL-PEADKLAIEWVSSQFRGTYRAGGELTK-----GREDMREQLDIVEROALIAQGGG 140
QY 116 ERIWTO-YFDROYTAS-----RAVAEVLNANGTEBPDGVEAFIDCEPIL 159
DB 141 TPAMTRLQGPQNPALPDLKALLAWOSKVTAVALRLKAFQSLDQPEDAF---DPI- 196
QY 160 RRRYFQVDEHR-SAEEOPLR-----MAPHYDISMTLLIQPTPCANGFVSLQAEVGG 210
DB 197 ---YSSBNHMKIVRYRGRDITGGDQGVGAHKDGGFLLTLQ---DNNKGLQVDYDG 248
QY 211 AFTDLPRPDVAVFCGATATLVGGQVAPRHHVAAPRRDQIAGSSRTSSVFLRPND 270
DB 249 SWVDVPIGTIVNIGELLEASNGYLAIVHVVTP---AGVERISVFFFSARLD 304
QY 271 FT-----FSVPLARECGFVSLDGETATFQDWIGGNYVNIIRTS 309
```

DB 305 ATIPILGSEELAQAARGPAS-DPDNPLFRD-VGTNVLKSRLRS 346

```
RESULT 6
US-09-924-841-5
; Sequence 5, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
; FILE REFERENCE: MS041-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; OTHER INFORMATION: from Streptomyces clavuligerus
; NAME/KEY: MUTAGEN
; LOCATION: (210)
; OTHER INFORMATION: Glu210 in native IPNS modified to Arg
US-09-924-841-5
```

Query Match 10.3%; Score 168; DB 9; Length 329;
Best Local Similarity 23.8%; Pred. No. 6e-09;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

```
QY 5 VPTSLAI-----QGLHODEFRCLDKGLFYLTDCGLDTBELKSAKDVIDPFHG 58
DB 10 VPTDISPLFTDAAAKRVAEIHGACGSGFFYATNHGV---DVQQLQDVNEFHGM 66
QY 59 SEAEK-----RAVTSVPVTRMGF-----TGLESESTAQITNTGSSDYSCMCSMGTAANLFPSSGPF 106
DB 67 TDQEKHDAIAHAYPNPDNHVANGYKAVPGKAVESFCYLNPDGDEHPRM-IAQTPHME 125
QY 107 DNLFPSSG---FERIWTQYFDROYTASRAVAEVLNANGTEBPDGVEA-----FLDCE 156
DB 126 VNLMPDEERHFRFPFCGGYRQMLKSTVLMRGIALALG-RPHFFDALAEODSLSSV 184
QY 157 PLLEPRYFQVPEHRSAAE-QLMAAPHYDISMTLLIQPTPCANGFVSLQAEVGAFTDL 215
DB 185 SLIRPYEERYPPVKTGPDGLSFRDHLDSMTITVLEQTQVON---LOVEYDGMWDI 240
QY 216 PYRPAVAVFCGATATLVGGQVAPRHHVAAPRRDQIAGSSRTSSVFLRPNDFTFSV 275
DB 241 PTSENDPLVNGCYTMAHTNDYFPAPNHRV-----KPVNAERLSLPFFLANGHEAVIE- 293
QY 276 PLARECGFVSLDGETATFQDWI 298
DB 294 PFVPE-GASEVRNEALSVDYL 315
```

```
RESULT 7
US-09-924-841-9
; Sequence 9, Application US/09924841
; Patent No. US20020127633A1
; GENERAL INFORMATION:
; APPLICANT: Dilley, David R
; APPLICANT: Kadyrzhanova, Dina K
; APPLICANT: Wang, Zhenyong
; APPLICANT: Warner, Toni M
; TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and
; TITLE OF INVENTION: Cephalosporins Under the Control of Bicarbonate
```

```
FILE REFERENCE: MSU41-453
; CURRENT APPLICATION NUMBER: US/09/924,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/413,231
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 9
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
; NAME/KEY: M01GEN
; LOCATION: (210)
; OTHER INFORMATION: GIu210 in native IPNS modified to Arg
US-09-924-841-9

Query Match          10.3%; Score 168; DB 9; Length 329;
Best Local Similarity 23.8%; Pred. No. 6e-09;
Matches 77; Conservative 50; Mismatches 150; Indels 46; Gaps 14;

OY 5 VPTFSLAEIQQGGLHODEFRRCRLDKGLFYLTDCGLDTTELKSAKDVIDFFEHG 58
DB 10 VPTIDISLPGFDAAKRAVAEIHGACRGSGFFPATMGV--DVQQLDVNVEFHGM 66
OY 59 SEAEK-----RAVTSVPVTRRGF---TGLESESTAQITGTSYSDYSCYMGST--A 106
DB 67 TDQEGHDLAIHAYNDDNPHVNRGKYKAVPGKRAVESFCYLNPDPGEDEHPM-IAAGTPWHE 125
OY 107 DNLPRPSGD----FERIMVQYFDROYTASRAVAEVLRAETGTEPDGVEA-----FLDCE 156
DB 126 VNLWDERRHPRFPFCGCTYRKMLKSTVLMRGIALALG-RPEHFPAALAEODSLSSV 184
OY 157 PLARFYPPOVPEHRSAAE-OPLRMAPHYDLSMTVLIQOTPCANGFVSLQAEVGAFPTDL 215
DB 185 SLIRFVYEYRPVVTGPDGQLSFRDLVDVSMITVLQTVQN---LQVETVDGMRDI 240
OY 216 PYRPDAVIVFCGAIATLVTVGGGVKAPRHVVAAPRRDQIAGSSRTSSVFFLPNADFTFSV 275
DB 241 PTESENDPVLNCGTYMAHTVNTDYPFAPNHRV-----KRVNAERLSLPFLNGCHEAVLE- 293
OY 276 PLARECGFVSLDGETATFODMI 298
DB 294 PFVPE-GASEEYRNALSYGDTL 315

RESULT 8
US-10-732-923-9964
; Sequence 9964, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 9964
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Rhodospseudomonas palustris
US-10-732-923-9964

Query Match          10.3%; Score 168; DB 17; Length 412;
Best Local Similarity 23.9%; Pred. No. 8.3e-09;
Matches 85; Conservative 39; Mismatches 141; Indels 90; Gaps 14;

OY 5 VPTFSLAEIQQGGLHODEFRRCRLD---KGLFYLTDCGLDTTELKSAKDVIDFFEHG 59
DB 10 VPTIDISLPGFDAAKRAVAEIHGACRGSGFFPATMGV--DVQQLDVNVEFHGM 66
```

```
DB 72 LPVLDSLRFRSDTAERAEFLRDVDAAFGRFFVYLVGHGISDRLIRDLVLFASRNFPAL-P 130
OY 60 EAEK---RAVTSVPVTRRGFTGLESESTAQITNTGTSYSDYSCYMGSTADNLFPSGDE 116
DB 131 EADKLDIEMINSP---HFRGYTPAGREFTR-----GQPDWEQLDVGAEREAFPLSRSA 181
OY 117 RIWTOYFDROYTASRAVAEVLRAETGTEPDGVEAFLDCEPLRFYFPQVPE----- 169
DB 182 PFWTR-----LOGPNQWPD---ALPELKPLI-LRYQDEVTELAKVLR 220
OY 170 -----HRSAEOPLRMAPHYDLSMTVLIQOTPCA 198
DB 221 VFAALQGAEDVPEPIYVSPNQLIKIRYPGRADAESDQGVTHKDSGFVITLLQDTVA 280
OY 199 NGVSVLQAEVGAFTDLPYRPDAVIVFCGAIATLVTVGGGVKAPRHVVAAPRRDQIAGSSR 258
DB 281 ---GLQVETADGMIDAPPLPGSFVNIGELIELANSGALRAVNRHVSPDPD---TVR 332
OY 259 TSSVFLRPVADFTFSV---PLARECGFVSLDGETATFODMIGGNVNIIRTS 309
DB 333 LSVAFFLGARLDATVTVLTPPELDAYARGVTODPQNPLFRD-VGRNVLKGRLR 386

RESULT 9
US-10-425-115-220323
; Sequence 220323, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 220323
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132519C.1.pep
US-10-425-115-220323

Query Match          9.5%; Score 156; DB 16; Length 342;
Best Local Similarity 23.4%; Pred. No. 1.2e-07;
Matches 78; Conservative 49; Mismatches 119; Indels 88; Gaps 16;

OY 4 TVPTFSLAEIQQGGLHODEFRRCRL---RDKGLFYLTDCGLDTTELKSAKDVIDFFEHG 59
DB 31 TSATVSLPIYDLSLGRDEVRAILBAGKEIGFQOVNNGVSLAQMOMETVCOEFTFR--L 88
OY 60 EAEKRAVTSVPVTRRGFTGLESESTAQIT-----NTGSYSDYSCYMGSTADNLF 111
DB 89 PAEDKA-----GLYSEDIGRAIRIYSMTFPGAARKYRDCIRLACS---PF 132
OY 112 S-GDPERIW-----TQYFDROYTASRAVAEVLRA-ATTEPD-----GGVEAF 152
DB 133 AVDSAAAMPDKRRRLREVEREFTVQTRGIMBILRLCEGLGRDLYEGDISGG--- 188
OY 153 LDCEPLLRFRYPPOVPEHRSAAEOPLRMAPHYDLSMTVLIQOTPCANGFVSLQAEVGAF 212
DB 189 ---DVLHVNHYPPCPDPNAT---LGLPFCRNLITLL-----LPSWPG-- 228
OY 213 TDLPYR-----PDVAVLFCGAIATLVTVGGGVKAPRHVVAAPRRDQIAGSSRTSVF 263
DB 229 LQVAVRSDWTRVPEVGAFFVNFGCQLEVTNGILKSIEHRVMTN-----LGVARTTVAT 283
OY 264 PLRPNADFTSVPLARECGFVSLDGETATFODM 297
DB 284 FIMPTTDLIG-PAAEFLSDNPNPCYRTLTFADF 316
```



```
Db      124 EKVMEP-----GGIGGHAFFVPSDDDKLDMCNMLALGVBPATIRRNMLPTTPAN 175
      115 FERLWTOYFROYTASRAVAREVLATGTEPDGGVEAFLDCEPLIRFPYQVBEHSAE 174
      176 FSKTLEKYSVEIRELCVRLLEHIAALGLAPRLMGVEAVQVRNMFPPCR-----P 231
      175 EQLPMAPHYLSWVTLIQOTPCANGFVSLQAEVGAFDLPYRPDAVLVFCGAIATLVT 234
      232 ELVLGLSHSDGSAVTVLQDDAFAFAGLOVLRG--GGGVAAVHPVPGALVNAVGGDTLEVL 289
      235 GGQVAPRRH-VAAFRDQIAGSSRTSSVFLRPVADTFESVPLARECGFDVSLDGETAT 293
      290 NGRKYSVHBRAVASEHD-----RMSVVTFAPAVDVLC-PLP-----ELVADGERR 337
      294 FQDMIGNYVNIIRTSK 310
      338 YRTYHGEYSRHHVTSR 354
```

RESULT 13

```
US-10-425-115-220322
; Sequence 220322, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 220322
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132518C.1.pcp
US-10-425-115-220322
```

```
Query Match      9.4%; Score 154; DB 16; Length 342;
Best Local Similarity 23.4%; Pred. No. 2e-07;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;
```

```
      4 TVPTFSIAELQGLHODEFRRLCL---RDKGLFYLTDCGLTDTLKSANDIVIDFEHGS 59
      31 TSAVTSLPIDLSIGRDEVRRAILEAGKEIGFQVNVHGSLAMQDMETVCGEFFF--L 88
      60 EAEKRAVTSVPVPMRGRGTGSESTAQIT-----NTGSYSDYSMCYMGTAADNLFP 111
      89 PAEDKA-----GLYSEDTGRATRIYSTMTFDGEEKYWRDCLRLACS---FP 132
      112 S-GDPERIW-----TOYFDRQYASRAVAREVLR---ATGTEPD-----GGVEAF 152
      133 AVGDSAAAMPDKRRLREVERFTYQTRGLGMEILRLCEGLGRPDVLEBDISGG----- 188
      153 LDCEPLIRFPYQVBEHSAEEOPLRMAPHYDLSWVTLIQOTPCANGFVSLQAEVGAF 212
      189 ---DVLVHNVHPCCDPNAT---LGLPPHCDRNLTLT-----LPSWVPG-- 228
      213 TDLPYR-----PDAVLVFCGAIATLVGGQVAKARHHVAAPRRDQIAGSSRTSSVF 263
      229 LEVAIRGDMIRVBPVGAFAVNFQGLLEVVTNGILKSIHRVMTN-----LGVARTIVAT 283
      264 FLRPVADTFPSVPLARECGFDVSLDGETATFQDM 297
      284 FIMPTTDLCLIG-PAAEFLSDNDNPPCYRTLITFGDF 316
```

RESULT 14

```
US-10-425-114-62424
; Sequence 62424, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62424
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3357-058-F2_FLI.pcp
US-10-425-114-62424
```

```
Query Match      9.4%; Score 154; DB 15; Length 371;
Best Local Similarity 23.4%; Pred. No. 2.3e-07;
Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;
```

```
      4 TVPTFSIAELQGLHODEFRRLCL---RDKGLFYLTDCGLTDTLKSANDIVIDFEHGS 59
      60 TSAVTSLPIDLSIGRDEVRRAILEAGKEIGFQVNVHGSLAMQDMETVCGEFFF--L 117
      60 EAEKRAVTSVPVPMRGRGTGSESTAQIT-----NTGSYSDYSMCYMGTAADNLFP 111
      118 PAEDKA-----GLYSEDTGRATRIYSTMTFDGEEKYWRDCLRLACS---FP 161
      112 S-GDPERIW-----TOYFDRQYASRAVAREVLR---ATGTEPD-----GGVEAF 152
      162 AVGDSAAAMPDKRRLREVERFTYQTRGLGMEILRLCEGLGRPDVLEBDISGG----- 217
      153 LDCEPLIRFPYQVBEHSAEEOPLRMAPHYDLSWVTLIQOTPCANGFVSLQAEVGAF 212
      218 ---DVLVHNVHPCCDPNAT---LGLPPHCDRNLTLT-----LPSWVPG-- 257
      213 TDLPYR-----PDAVLVFCGAIATLVGGQVAKARHHVAAPRRDQIAGSSRTSSVF 263
      258 LEVAIRGDMIRVBPVGAFAVNFQGLLEVVTNGILKSIHRVMTN-----LGVARTIVAT 312
      264 FLRPVADTFPSVPLARECGFDVSLDGETATFQDM 297
      313 FIMPTTDLCLIG-PAAEFLSDNDNPPCYRTLITFGDF 345
```

```
RESULT 15
US-10-425-114-53454
; Sequence 53454, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53454
; LENGTH: 372
; TYPE: PRT
```

ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700169504_FLI.pep
US-10-425-114-53454

Query Match 9.4%; Score 154; DB 15; Length 372;

Best Local Similarity 23.4%; Pred. No. 2.3e-07;

Matches 78; Conservative 48; Mismatches 120; Indels 88; Gaps 16;

```
Qy 4 TVPTSLAELOQGLHODEFRCL---RDKGLPYLTDCGLTDTLKSADIVDPFHGS 59
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 TSATVSLPIVDLSLGRDEVRAILLEAGKEIGFFQVNHGVSLEAMQDMETVCQEFFR--L 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 60 EAKKRAVTSPPVPTMRGFTGLESESTAOIT-----NTGSYSDYSCYSGTADNLP 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 119 PAEDKA-----GLYSBDTGRATRIYSSMTFDTGGEKTYRDCRLACS--FP 162
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 112 S-GDFERIW-----TOYFDRQYTA$RAVAREVL$---ATGTEPD-----G$VEAF 152
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 163 AVGDSAAAMPDKPRRLREVERFTVQTRGLGMEILRLCEGLGLRPDYLEGDISGS---- 218
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 153 LDCEPLRRRYFPQVPEHRSABEQPLRNAPHYDLSNVTLIQOTPCANGFVSLQABVGAF 212
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 219 ---DVLHVNHYPPCPDPNAT---LGLPHCDRNILTL-----LPSWVPG-- 258
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 213 TDLPYR-----PDATLVFCGAIATLVYTGGOVKAPRRHHVAPRRDQIAGSRTSSVF 263
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 LEAVIRGDMIRVEPVPGAFVNVFPGCOLEVTNGILKSIHRVMTN-----LGVARTVAT 313
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 264 FLRPNADFTFSVPLARECGFDVSLDGTATFQDW 297
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 314 FIMPTTDCIG-PAABFLSDDNPPCYRTITFGDF 346
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: September 19, 2005, 15:24:23
Job time : 86 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 19, 2005, 15:19:28; Search time 4917 Seconds

(without alignments)
3064.790 Million cell updates/sec

Title: 10719236-1_232-1164

Perfect score: 1636

Sequence: 1 MDTVPTFTSLAELOGLQHD.....ATFQDWIGNVNIRRTSKA 311

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cgn2.1/USPTO_spool/h/10719236-1/TRANS/runat_19092005_153157_7740/app_query.fasta_1.45
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=pcr -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=10719236-1/TRANS/@CGN_1_1_3731_@runat_19092005_153157_7740 -NCPU=6
-ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_da:.*
2: gb_hlg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_srs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1636	100.0	1230	6	107104 Sequence 1
2	1636	100.0	1230	6	AR479090 Sequence
3	1634	99.9	936	6	AX300747 Sequence
4	1634	99.9	2494	1	STWCERDA M32324 S.clavulig

5	1385	84.7	936	1	AY318742
6	1378	84.2	1141	1	AF317908
7	1300	79.5	936	1	AY318743
8	1210	74.0	939	6	A42999
9	1210	74.0	939	6	A72870
10	1210	74.0	939	6	AR023763
11	1195.5	73.1	942	6	A43001
12	1195.5	73.1	942	6	A72872
13	1195.5	73.1	942	6	AR023762
14	1195.5	73.1	942	6	191795
15	1195.5	73.1	942	6	NLCERF
16	1195.5	73.1	2169	1	SS5763
17	1195.5	73.1	939	1	AY318744
18	1012	61.9	933	1	NLCERF
19	972.5	59.4	933	1	STWCERF
20	946.5	57.9	954	1	STWCERF
21	944	57.7	6358	1	LDYK90C
22	933.5	57.1	1511	6	105489
23	933.5	57.1	4623	6	CQ796893
24	924.5	56.5	5775	8	ACH404737
25	427	26.1	5960	1	SCCERGEN
26	183	11.2	1044	1	STWIPNSSJ
27	180	11.0	346357	1	SGPCBC
28	178	10.9	990	1	AP003010
29	169	10.3	340857	1	STWIPNS
30	168	10.3	990	1	AT01132
31	168	10.3	990	6	AT763606
32	168	10.3	29870	6	AX572600
33	168	10.3	349640	1	AB016866
34	166.5	10.2	310029	1	AF320779
35	165	10.1	990	1	AR029451
36	165	10.1	1250	6	FRPCBC
37	164	10.0	1310	1	AB045851
38	163	10.0	786	1	AY260760
39	161	9.8	36687	1	AB045853
40	160	9.8	786	1	AB045852
41	158	9.7	786	1	AB045857
42	158	9.7	786	1	BR013114
43	158	9.7	1331	8	AT075378
44	156.5	9.6	1891	3	AY102699
45	156.5	9.6	1938	3	AY102699

ALIGNMENTS

RESULT 1

LOCUS	107104	1230 bp	DNA	linear	PAT 02-DEC-1994
DEFINITION	Sequence 1 from Patent EP 0341892.				
ACCESSION	107104				
VERSION	107104.1	GI:589821			

KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 1230)

AUTHORS
Ingolia,T.D., Kovacevic,S., Miller,J.R. and Skatrud,P.L.

TITLE
Recombinant DNA expression vectors and DNA compounds that encode

JOURNAL
deacetoxycephalosporin C synthetase

FEATURES
Patent: EP 0341892-A1 1 15-NOV-1989;
location/Qualifiers
1..1230
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.:	2.7e-136	Length:	1230
Score:	1636.00	Matches:	311
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

10719236-1_232-1164 (1-311) x 107104 (1-1230)

QY 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
Db ATGACACGACGAGCGTGCCACCTTCAGCTGGCCGCAACTCCGACGAGGCGCTGCACCAAGAC 291

QY 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
Db GAGTTCGCGAGGTGTCTGAGGGAACAAGGCGCTCTTCTATCTGACGAGCTGGCGTCTGACC 351

QY 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGlnHisGlySerGlu 60
Db GACACCGAGCTGAGTGGCCGCAAGACATCTCATCTTCCTTCGAGCAGCGACGAGGAG 411

QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db GCGGAGAAAGCGCGCGTCACTCCCGCCGACCAATGCGCGCGGCTTCACCGGCGTGC 471

QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db GAGTCGAGAGACCGCGCCAGATCAACAATACCGGAGCTACTCCGACTCATGTGC 531

QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
Db TACTCGATGGGCAACCGGGAACAACCTCTCCGTCGGTGACTTCGAGCGGATCTGAGAC 591

QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db CAGTACTTCGACCGCCAGTACACCGCCCTCCGCGCGGTCCGCGGAGGTCTTCGCGG 651

QY 141 ThrGlyThrGluProAspGlyGlyValGluValaPheLeuAspCysGluProLeuLeuArg 160
Db ACCGGAGCCGAGCCCGAGCGGCGGTGAGGCTTCTTCGACTCGAGCGCTCTGCGG 711

QY 161 PheArgTyrPheProGlnValProGlnHisArgSerAlaGluGlnGlnProLeuArgMet 180
Db TTCGCTACTCTCCGCGAGGTCCCGAGGACCGGACCGCGGAGGACGCGCTCGGATG 771

QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db GCGCGGACATACGACCTGTGATGTGACCTCTCATCCGACGACACCTTACCTCGGAC 831

QY 201 PheValSerLeuGlnAlaGluValGlyValaPheThrAspLeuProTyrArgProAsp 220
Db TTCGTACGCTTCAGGCGGAGGTGCGCGCGCGCTTCACGAGCTTCCCTTACCTCGGAC 891

QY 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLysAla 240
Db GCCGTCCGTCTCTGCGGCGCATGCGGACCTGGTGACCGGCGGCGGAGGTCAAGGCC 951

QY 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db CCCCCGACCACTGTGCGCGCGCCCGCGGAGCAAGATAGCGGCGGAGCGGACGCACTCC 1011

QY 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db AGTGTGTTCTTCTCTCTCCCAAGCGGACTTCACTTCTCTCGTCCGCTGCGCGGAG 1071

QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGlyGly 300
Db TGGCGGCTTCGATGTGACGCTGAGCGGAGACCGGCAAGTTCAGAGATTGATCGGGGCG 1131

QY 301 AsnTyrValaIleHisArgArgThrSerLysAla 311
Db AACTACGTGAACATCCCGCGACATCCAAAGGCA 1164

RESULT 2
LOCUS AR479090 1230 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 1 from patent US 6699699.
ACCESSION AR479090
VERSION AR479090.1 GI:47238046

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1230)
AUTHORS Yunn-Bor,Y., Chia-Li,W., Jyh-Shing,H. and Ying-Chieh,T.
TITLE Mutated penicillin expandases
JOURNAL Patent: US 6699699-A 1 02-MAR-2004;
FEATURES
source
Location/Qualifiers
1..1230
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2,7e-136 Length: 1230
Score: 1636.00 Matches: 311
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

10719236-1_232-1164 (1-311) x AR479090 (1-1230)

QY 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
Db ATGACACGACGAGCGTGCCACCTTCAGCTGGCCGCAACTCCGACGAGGCGCTGCACCAAGAC 291

QY 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
Db GAGTTCGCGAGGTGTCTGAGGGAACAAGGCGCTCTTCTATCTGACGAGCTGGCGTCTGACC 351

QY 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGlnHisGlySerGlu 60
Db GACACCGAGCTGAGTGGCCGCAAGACATCTCATCTTCCTTCGAGCAGCGACGAGGAG 411

QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db GCGGAGAAAGCGCGCGTCACTCCCGCCGACCAATGCGCGCGGCTTCACCGGCGTGC 471

QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db GAGTTCGCGAGGTGTCTGAGGGAACAAGGCGCTCTTCTATCTGACGAGCTGGCGTCTGACC 531

QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
Db TACTCGATGGGCAACCGGGAACAACCTCTCCGTCGGTGACTTCGAGCGGATCTGAGAC 591

QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db CAGTACTTCGACCGCCAGTACACCGCCCTCCGCGCGGTCCGCGGAGGTCTTCGCGG 651

QY 141 ThrGlyThrGluProAspGlyGlyValGluValaPheLeuAspCysGluProLeuLeuArg 160
Db ACCGGAGCCGAGCCCGAGCGGCGGTGAGGCTTCTTCGACTCGAGCGGCGCTGCGCG 711

QY 161 PheArgTyrPheProGlnValProGlnHisArgSerAlaGluGlnGlnProLeuArgMet 180
Db TTCGCTACTCTCCGCGAGGTCCCGAGGACCGGAGCGCGGAGGACGCGCTCGGATG 771

QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db GCGCGGACATACGACCTGTGATGTGACCTCTCATCCGAGACACCTTCCGCGCAAGGCC 831

QY 201 PheValSerLeuGlnAlaGluValGlyValaPheThrAspLeuProTyrArgProAsp 220
Db TTCGTACGCTTCAGGCGGAGGTGCGCGCGCGCTTCACGAGCACTGCGCTTCCGAGC 891

QY 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLysAla 240
Db GCCGTCCGTCTCTCTCTCCCAAGCGGACTTCACTTCTCTCGTCCGCTGCGCGGAG 951

QY 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260

Db 952 CCCCGCACCATGTGCGGCGCCCCCGGACGAGCACGATGCGGGGACGAGCGCGACCTCC 1011
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db 1012 AGTGTGTTCTTCTCCGTCCCAACGGGAGCTTCTCCGTCCGCTGCGGCGCGAG 1071
Qy 281 CysGlyPheAspValSerLeuAspGlyValThrAlaThrPheGlnAspTrpIleGlyGly 300
Db 1072 TCCGCTTCGATGTCAGCTTGACGCGGACCGACCGACCTTCAGGATTTGATCGGAGGC 1131
Qy 301 AsnTrValAsnIleArgArgThrSerIysAla 311
Db 1132 AACTACGTGAACATCGCGGACATCCAGGCA 1164

RESULT 3
AX300747 936 bp DNA linear PAT 30-NOV-2001
LOCUS Sequence 1 from Patent WO185951.
DEFINITION AX300747
ACCESSION AX300747
VERSION AX300747.1 GI:17382047
KEYWORDS
SOURCE Streptomyces clavuligerus
ORGANISM Streptomyces clavuligerus
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycinae; Streptomycetaceae; Streptomyces.
REFERENCE 1 Johnson, R.I. and Newbert, R.W.
AUTHORS A modified expandase and uses thereof
TITLE Patent: WO 0185951-A 1 15-NOV-2001;
JOURNAL ACS Dobfar UK Limited (GB)
FEATURES
source location/Qualifiers
1..936
/organism="Streptomyces clavuligerus"
/mol_type="unassigned DNA"
/db_xref="taxon:1901"
1..936
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD13038.1"
/db_xref="GI:17382048"
/translation="MDTVPFPLAELOOGLHODEFRRLRDKGLFVLTDCGLDPTTEL
KSAKDLVIDFPHSGSEARKRAVTSVPMTREGTGLSESTAOITNCSYSDYSCYS
MGTDNLTPSGDFERITQYPRDQYTSRAVAEVLKATGTBPDCGVAFPLDCEILR
FRYPQVBEHSAEOPLEMAHPHDLNVTLLIQTPCANGFVSLQAEVGAFTDLPYR
PDVALVFCGAIATLVTGQVNAPRHVAAPRDQIAGSSRTSSVFELRPNDFTFSVP
LARECGFVSLDGETATFQDMTIGNYNIRTSKA"

ORIGIN
Alignment Scores: 2.97e-136 Length: 936
Pred. No.: 1634.00 Matches: 310
Score: 100.00% Conservative: 1
Percent Similarity: 99.68% Mismatches: 0
Best Local Similarity: 99.88% Indels: 0
Query Match: 6 Gaps: 0
10719236-1_232-1164 (1-311) x AX300747 (1-936)

Qy 1 MetAspTrpThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAsp 20
Db 1 ATGGACACGACGCTGCCACCTTCAGCTTGCGGCAACTCCAGAGGGCTGACACGAGAC 60
Qy 21 GluPheArgArgCysLeuArgAspIysGlyLeuPheTrpLeuThrAspCysGlyLeuThr 40
Db 61 GAGTTCGCCAGGTGCTCTGAGGACAAAGGCGCTCTTCTATCTGACGAGACTGCGGTCTGACC 120
Qy 41 AspTrpGluLeuLysSerAlaIysAspIleValIleAspPhePheGlnHisGlySerGlu 60
Db 121 GACACCGAGCTTAAGTCGGCCAAAGACCTCGTCATCGACTTCTTGGAGACGGACGAG 180
Qy 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80

Db 181 GCGAGAAAGCGCGCGTCACTCGCCGTCCCGACCAACAGCGCGCGGCTTACCGGGCTG 240
Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTrpSerAspTrpSerMetCys 100
Db 241 GAGTCGGAGACACCGCCAGATCAACAAATACCGGAGCTACTCCGACTACGTGATGAC 300
Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluAspIleThrThr 120
Db 301 TACTGATGGGACCGCGGACCAACTCTCCGTCGGGTGAGTACTGACCGGATCTGAGACC 360
Qy 121 GlnTrpPheAspArgGlnTrpThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db 361 CAGTACTTCGACCGCGAGTACACCGCTCCCGCGGGTCTGCGCGGAGGCTCTGGGGCG 420
Qy 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db 421 ACCGGACCGAGCCCGAGCGGGGTGAGGCTCTTCTCGACTGCGACCCGCTGCGCG 480
Qy 161 PheArgTrpPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMet 180
Db 481 TTCGCTACTTCCCGAGGTCCCGAGCACCGAGCGCGAGAGCAGCCCTGCGAGTG 540
Qy 181 AlaProHisTrpAspLeuSerMetValThrLeuIleGlnGlnTrpProCysAlaAsnGly 200
Db 541 GCGCGGACCTAGACACTGTGATGCTACCTTCATCCAGAGACACCTGCGGCCACAGCG 600
Qy 201 PheValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeuProTrpArgProAsp 220
Db 601 TTCGTCAGCTCCCGAGCGCGAGGTGCGCGCGCTGTCAGGACCTGCTTACCGTCCGAC 660
Qy 221 AlaValLeuValPheCysGlyValAlaIleAlaThrLeuValThrGlyGlyValIysAla 240
Db 661 GCGTCTCTGCTTCTGCGCGCCATCGCACTGTGACCGGCGCGGCGGAGGTCAAGGCC 720
Qy 241 ProArgHisIysValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 721 CCCCGGACCATGTGCGGGCCCGCCCGGAGGACCGAGATGAGGGGAGGACGCGACCTCC 780
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db 781 AGTGTGTTCTTCTCCGCTCCCAACGCGGACTTCACTTCTCCGCTCCGCTGCGGCGGAG 840
Qy 281 CysGlyPheAspValSerLeuAspGlyValThrAlaThrPheGlnAspTrpIleGlyGly 300
Db 841 TCCGCTTCGATGTCAGCTGAGCTGAGCGGACGCGCACGCTTCAGGATTTGATCGGAGGC 900
Qy 301 AsnTrValAsnIleArgArgThrSerIysAla 311
Db 901 AACTACGTGAACATCGCGGACATCCAGGCA 933

RESULT 4
STWCFDA 2494 bp DNA linear BCT 26-APR-1993
LOCUS S.clavuligerus isopenicillin N epimerase (cefi) gene and
DEFINITION deacetoxycephalosporin C synthetase (DMOCs) gene, complete cds.
ACCESSION M32324.1 GI:153203
VERSION M32324.1 M24140
KEYWORDS deacetoxycephalosporin C synthetase; isopenicillin N epimerase.
SOURCE Streptomyces clavuligerus
ORGANISM Streptomyces clavuligerus
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycinae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1520 to 2494)
AUTHORS Kovacevic,S., Weigel,B.J., Tobin,M.B., Ingolia,T.D. and Miller,J.R.
TITLE Cloning, characterization, and expression in Escherichia coli of
the Streptomyces clavuligerus gene encoding deacetoxycephalosporin
C synthetase
JOURNAL J. Bacteriol. 171 (2), 754-760 (1989)
MEDLINE 89123150
PUBMED 2644235
REFERENCE 2 (bases 1 to 1706)
AUTHORS Kovacevic,S., Tobin,M.B. and Miller,J.R.

TITLE The beta-lactam biosynthesis genes for isopenicillin N epimerase and deacetoxycephalosporin C synthetase are expressed from a single transcript in *Streptomyces clavuligerus*

JOURNAL J. Bacteriol. 172 (7), 3952-3958 (1990)

MEDLINE 90299822

PUBMED 16945525

COMMENT Original source text: S.clavuligerus DNA, clone POW80. Draft entry and computer-readable sequence for [1] kindly submitted by J.R.Miller, 26-FEB-1990, for release after publication.

FEATURES

source

1. 2494

/organism="Streptomyces clavuligerus"

/mol_type="genomic DNA"

/db_xref="taxon:1901"

152..21706

/product="cefD mRNA"

284..1480

/note="isopenicillin N epimerase (cefD)"

/codon_start=1

/transl_table=11

/protein_id="AA26714.1"

/db_xref="GI:153204"

/translation="MAVADWEARGRMILDPVVNLNTGSGPLPRSAFERVTGFRALIAEPMDFLREVPALIMQARESLARLIGDPLRLATLVTAIVTAIVASRLAEGELISDSEYPMRWCWERRVARRHGLRTRPLRPLDPAEITAAVAAMPRTLRPFPSHYVSTTGILPAAELCEERARGITTVVDGAAAPGFDLDSRIFCDPYASGHKMLAPTVGFPHLAPRLBELPEPTOVSMAYEPBGSGPAAADRGTGTPGLRKEGTRDICPWLATPESIDFOAELGPAARRARRLTDHARRLLADRPRTLLTPDPSPLSGMVAVRLPGTDAAELRGLMERFRLEAAVAEQPPVLIRISANFYETTEEIDRLADALDALTGE"

1559..2494

/note="deacetoxycephalosporin C synthetase (DAOCS)"

/codon_start=1

/transl_table=11

/protein_id="AA26715.1"

/db_xref="GI:153205"

/translation="MDTVPPTFSLAELQOGLHDEFFRCLEDKGLFYLTQGLTDTLSKASDLYIDFFHGSAEKRAVTSVPVPMRWGFTGLSESTAOITNTGYSIDSMCVSMGTADNLPSPSDPERITQYDPROYTAARARVLAATGEPGCGIAPFLDCEPLLRFRYPQYPERHSAEQPLRKMAPHYDLSKVTLLIQTPCANGVSLQAEVGGATDLPKRPDAVLCGAIATLVGGQVAPRHVAAPRRDQIAGSSRTSSVFLIRPADPFTSVPLARECGFVSLDGETATFQDWIGGNVNIIRTSKA"

ORIGIN

Alignment Scores:

Pred. No.: 9.26e-136 Length: 2494

Score: 1634.00 Matches: 310

Percent Similarity: 100.00% Conservative: 1

Best local Similarity: 99.68% Mismatches: 0

Query Match: 99.88% Indels: 0

DB: 1 Gaps: 0

10719236-1_232-1164 (1-311) x STMCFEDA (1-2494)

Qy 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAsp 20

Db 1559 ATGACACGACGCGTCCACCTTCAGCCTGCGCCAGACTCCAGCAGGCGCTGCACAGGAC 1618

Qy 21 GluPheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40

Db 1619 GAGTTCGCGAGGTGTCTGAGGAGCAAGGCGCTCTCTATCTGACGAGACTGCGGTCTGACC 1678

Qy 41 AspThrGluLeuValSerAlaIysAspIleValIleAspPhePheGlnHisGlySerGlu 60

Db 1679 GACACCGAGCTGAGACTCGCCAGCAAGCACTCGTCATCGACTTCTGACACGAGCAGCAG 1738

Qy 61 AlaGluValArgAlaValThrSerProValProThrMetCysArgGlyPheThrGlyLeu 80

Db 1739 GCGGAGAAAGCGCGCGTCACTCGCCGCTCCACCAAGCGCGCGGCGCTTACACCGGGCTG 1798

Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100

Db 1799 GAGTTCGAGAGACACCGCCAGATCACCAATACCGGCACTACTCGACTACGATGTGC 1858

Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTrpThr 120

Db 1859 TACTCGATGGGACCGCGGACCAACTTCTCCCTCGGTGACTTCCAGCGGATCTGGACC 1918

Qy 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140

Db 1919 CAGTACTTCGACCGCCAGTACACCGCTCCCGGCGGTGCGCCGGAGAGTCTCGGGGCG 1978

Qy 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160

Db 1979 ACCGGGACCCAGCGCCGACGCGGCGGTGAGAGCCCTTCTTCATCGACGCGCGCTCGCG 2038

Qy 161 PheArgTyrPheProGlnValProGluHisArgSerAlaGluGlnInProLeuArgMet 180

Db 2039 TTCCGCTACTTCCCGAGAGTCCCGAGACACCGAGCGCCAGAGACAGCCCTCGCGATG 2098

Qy 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200

Db 2099 GCGCGGACACTACGACCTGTGATGTACCTCATCCAGCAGACACCTCGCGCAACGGC 2158

Qy 201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220

Db 2159 TTGCTCAGCTCCAGGCGCGAGTCTGGCGGCGGCTTCACGAGCCTGCGCTACCGCGGAC 2218

Qy 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLysAla 240

Db 2219 GCGGCTCTGCTCTTCTGCGGCGCATCGACCTGTGTGACCGCGCGCCAGGTCAAGGCC 2278

Qy 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260

Db 2279 CCCCGGACCATGTGCGGCGCCCCCGAGGAGCAAGTACGGGACGACCGCACCTCC 2338

Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280

Db 2339 AGGTGTTCTTCTTCCTCGTCCCAACGGGACTTACCTTCCCTCCGCTGCGCGCGGAG 2398

Qy 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300

Db 2399 TCGCGCTTCATGTACCTGACCTGACGCGGAGACCGCCACGTTCCAGGATTTGATCGGGGC 2458

Qy 301 AsnTyrValAsnIleArgArgThrSerTyrAla 311

Db 2459 AACTAGCTGAACATCGCGCCACATCCAGGCA 2491

RESULT 5

AY318742 936 bp DNA linear BCT 06-OCT-2004

LOCUS

DEFINITION Streptomyces ambofaciens deacetoxycephalosporin C synthase gene, complete cds.

ACCESSION AY318742

VERSION AY318742.1 GI:32492591

SOURCE

ORGANISM Streptomyces ambofaciens

REFERENCE Streptomyces ambofaciens Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

AUTHORS 1 (bases 1 to 936)

TITLE Huu,J.S., Yang,Y.B., Deng,C.H., Wei,C.L., Liaw,S.H. and Tsai,Y.C.

JOURNAL Family shuffling of expandase genes to enhance substrate specificity for penicillin G

PUBMED Appl. Environ. Microbiol. 70 (10), 6257-6263 (2004)

REFERENCE 15466573

AUTHORS 2 (bases 1 to 936)

TITLE Huu,J.S., Yang,Y.B., Wei,C.L. and Tsai,Y.C.

JOURNAL Cloning the DAOCS gene from Streptomyces ambofaciens

PUBMED Unpublished

AUTHORS 3 (bases 1 to 936)

TITLE Huu,J.S., Yang,Y.B., Wei,C.L. and Tsai,Y.C.

JOURNAL Direct Submission

PUBMED Submitted (08-JUN-2003) National Yang-Ming University, Institute of Biochemistry, 115, Li-Nong St. Sec. 2, Shih-Pai, Taipei 112, ROC

FEATURES Location/Qualifiers

```
source
1. .936
/organism="Streptomyces ambofaciens"
/mol_type="genomic DNA"
/db_xref="taxon:1889"
CDS
1. .936
/feature="DAOCs"
/codon_start=1
/translation="deacetoxycephalosporin C synthase"
/protein_id="AAP85380.1"
/db_xref="GI:32492592"
/translation="MDTVPTFTHLAELQELHODEFRSCLAREKVEFYLTGSSLEADQ
KSAQDVVDFPEHGEERAAVTSVPVTRRGTGSESTAOITNTGYSYMWCS
MGTATNPTLADPERIWMYHPRMWDVAEAPRQVTKATGREGVVAFWNCBPLR
FRYFPEVPAHRSABEVPFLMAHYDSTVTVIOQTPCPNGPSVLOCVEGCTVYDLPAK
PAVALVFCGATVLTGGOVKAPRKHVAAPGRDRIAGSSRTSSVFPLRPSDFSPSV
LARGCFDILSDGTAFFKDWIIEGVNVIIRKSEA"
```

ORIGIN

Alignment Scores:

Pred. No.:	4,62e-114	Length:	936
Score:	1385.00	Matches:	253
Percent Similarity:	91.96%	Conservative:	33
Best Local Similarity:	81.35%	Mismatches:	25
Query Match:	84.66%	Indels:	0
DB:	1	Gaps:	0

10719236-1_232-1164 (1-311) x AY318742 (1-936)

```
QY 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
Db 1 ATGGACACACGCGTACCGAAGTTCACCTCGCGGACCTTCAGAGGGGCTCCACCGAGAC 60
QY 21 GluPheArgArgCysLeuArgAspGlyLeuPheTyrlleuThrAspCysGlyLeuThr 40
Db 61 GAGTTCGGTAGCTGTCTGAGGAGGAGGCGGCTTTCTATCTGACCGCGCAGCTGCTGCG 120
QY 41 AspThrGluLeuSerLeuAlaValAlaAspPheGluHisGlySerGlu 60
Db 121 GAGCGCGACGAGAAATCCCGCAAGATGTCGTCGATTTCTTCAGACCGGTACGGAG 180
QY 61 AlaGluValArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db 181 GAGGAGAACGGGGCCCTCATCATCCGCTTCCACCATCCGCGCGGATTCACCGGGCTG 240
QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrlSerAspTyrSerMetCys 100
Db 241 GAGTCGAGAGACACCGCCAGATCACCAACACCGGACCTACTCCGACTACGATGTCG 300
QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
Db 301 TACTCGATGGGACCCCGGCAATATATTTCCACCGGACTCTTCAGACGATCTGGAGG 360
QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db 361 CACTACTTGAACCGATATGATGACGCCACCGGGAAGCGGACGAGCTCTGAAGGGG 420
QY 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db 421 ACCGGGACCGGAACCGGAGCGCGCTCGAAGACTTCGTAACCTGTGAAACCGCTGCGAG 480
QY 161 PheArgTyrPheProGlnValProGlnHisArgSerLeuAlaGluGlnProLeuArgMet 180
Db 481 TTCCCGGATTTCCCGGAGTCCCGGCGCATCGAGGGCCGAGAGAGTGCTCTGGCGATG 540
QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db 541 GGGCGGCAATTAGATTTGTGCAACGATGAGCGTGAATTCACACAGACCGCGCGGAAGCG 600
QY 201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db 601 TTGTGAGCTCTGACAGTGCAGGTCGGCGGACCTTTGTGACCTCGCGCAACCGCGAT 660
```

```
QY 221 AlaValLeuValPheCysGlyValAlaIleAlaThrLeuValThrGlyGlyValLeuVal 240
Db 661 GCGGTTCTTGCTTCTTGGCAGACCGCTCGGACCTCGGTGACCGGCGGCGGATGAAGGCC 720
QY 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 721 CCGAAGCATCATCTCGCGGCGCGCGGCGGAGCGGATTCGAGGACGAGCGCGTACGTCC 780
QY 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db 781 ACTGCTCTTCTTCTCGCGCGCGCATCTCGACCTTCTCGGTCCGCTGCGCGAGCGAG 840
QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspThrIleGlyGly 300
Db 841 TCGGATTCGACATCAGCTCGGACCGGACCGACCGGACGATTCAGGACTGTGACAGGCG 900
QY 301 AsnTyrValAsnIleArgArgThrSerLysAla 311
Db 901 AACTACGTCAATATCCGAGGAATCCGAGGCA 933
```

RESULT 6
AF317908 1141 bp DNA linear BCT 02-OCT-2001
LOCUS
DEFINITION Streptomyces jumonjensis deacetoxycephalosporin C synthase gene,
complete cds.
ACCESSION AF317908
VERSION AF317908.1 GI:15824728
KEYWORDS

SOURCE

Streptomyces jumonjensis
Streptomyces jumonjensis
Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
Streptomyces; Streptomycetaceae; Streptomyces.

REFERENCE

1 (bases 1 to 1141)
Sim.T.S. and Sim.J.
Cloning and purification of Streptomyces jumonjensis
deacetoxycephalosporin C synthase

JOURNAL

Unpublished
2 (bases 1 to 1141)
Sim.T.S. and Sim.J.
Direct Submission
Submitted (01-NOV-2000) Microbiology, National University of
Singapore, 5 Science Drive 2, Singapore 117597, Singapore

FEATURES

location/Qualifiers
1. 1141
/organism="Streptomyces jumonjensis"
/mol_type="genomic DNA"
/strain="ATCC 29864; NRRL 5741"
/db_xref="ATCC:29864"
/db_xref="taxon:1945"

CDS

38..973
/note="DAOCs; oxidoreductase; involved in antibiotic
biosynthesis; iron and alpha-ketoglutarate required for
enzyme activity"
/codon_start=1
/translation="MDTVPTFTHLAELQELHODEFRSCLAREKVEFYLTGSSLEADQ
KSAQDVVDFPEHGEERAAVTSVPVTRRGTGSESTAOITNTGYSYMWCS
MGTATNPTLADPERIWMYHPRMWDVAEAPRQVTKATGREGVVAFWNCBPLR
FRYFPEVPAHRSABEVPFLMAHYDSTVTVIOQTPCPNGPSVLOCVEGCTVYDLPAK
PAVALVFCGATVLTGGOVKAPRKHVAAPGRDRIAGSSRTSSVFPLRPSDFSPSV
LARGCFDILSDGTAFFKDWIIEGVNVIIRKSEA"

ORIGIN

Alignment Scores:

Pred. No.:	2,45e-113	Length:	1141
Score:	1378.00	Matches:	254
Percent Similarity:	90.68%	Conservative:	28
Best Local Similarity:	81.67%	Mismatches:	29
Query Match:	84.23%	Indels:	0
DB:	1	Gaps:	0

10719236-1_232-1164 (1-311) x AY317908 (1-1141)

QY 1 MetAapThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
DB ATGACACGACGCGTCCGACGCTTACGCTGGCCGAGCTTCCAGAGGGGCTCCACACGAG 97

QY 21 GluPheArgArgCysLeuArgAspIlysglyLeuPheTyrlleuThraAspCysGlyleuThr 40
DB GAGTCCCGAGCTGCTGCGGAGACAGGTCCTCTTCATCTGACGAGACGCGGCTGTGC 157

QY 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPheGlnHisGlySerGlu 60
DB GAGCCCGCATCAGAAATCCGCGCAAGACGTCGACATTGACTTCTTGCAGACATGCGACG 217

QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
DB GAGGAGAGCGGGCGCGCACCTCCACGATTCCTCCACATTCGCGGGGCTTCACGGGGTTG 277

QY 81 GluSerGluSerThrAlaGlnIleThraPheThrGlySerTyrlleuAspTyrlleuSerMetCys 100
DB GAGTGGAGAGACCGACAGATCACACCGCGGACCTACTCCGACTCTCGATGATGTC 337

QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
DB TACTCGATGGGCTGGCGGACAGATCTTCCCGCGGGGACTTTCAGACGGGCTGTGACG 397

QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
DB CACTATTTCCGCGCGGTATGATGACCTCACCCAGAGAGTGGCGCGGCTCTGAGACG 457

QY 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGlyProLeuLeuArg 160
DB ACCGAGACCGAACCGGTCGGCGCGCTCGAGTCTTGTGACGTGCGAACCCCTGCTGAG 517

QY 161 PheArgTyrPheProGlnValProGlnHisArgSerAlaGluGluGlnProLeuArgMet 180
DB TTCGGTACTTCCCGAGAGTCCCGAGAGCGCACGCGCGAGACAGCCCTCGCGGAG 577

QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
DB GCGCGCGACTACGACCTGTGATGTGACGAGTACGACACCCGTCGCGGACCGG 637

QY 201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
DB TTCGTCAGCTTCCAGGCGGAGATCGGCGGCTTCTGTGACCTCCGCGCACGCGGAC 697

QY 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLysAla 240
DB GCGGTCTCTGCTTCTGCGGAGCCATCGGACCTGTGTGACCGGCGCGCAGGTGAAAGCC 757

QY 241 ProAlaGlnHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
DB CCGAGGACCATAGTGGCGCGCGCGGACCGGAGCCAGATCGCGGACGCGCGCTGCTCC 817

QY 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
DB AGCGCTTCTTCTTCTCGCGCCGACCTTCACTTCTCGGTCTTCTTCTTCTTCTTCTTCT 877

QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGlyGly 300
DB TCGCGATTCGATATACGCTGTGACGCGGAGACCGCACGTTCAAGACACTGATGAGGCG 937

QY 301 AsnTyrValAsnIleArgArgThrSerLysAla 311
DB AACTATGTGAACATTCGACGAGACGTCGAGCA 970

RESULT 7
LOCUS AY318743 936 bp DNA linear BCT 06-OCT-2004
DEFINITION Streptomyces chartreusis deacetoxycephalosporin C synthase gene,
complete cds.
ACCESSION AY318743

VERSION AY318743.1 GI:32492593
KEYWORDS
SOURCE Streptomyces chartreusis
ORGANISM Streptomyces chartreusis
REFERENCE Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
AUTHORS Heu, J.S., Yang, Y.B., Deng, C.H., Wei, C.L., Liaw, S.H. and Tsai, Y.C.
TITLE Family shuffling of expandase genes to enhance substrate
specificity for penicillin G
JOURNAL Appl. Environ. Microbiol. 70 (10), 6257-6263 (2004)
PUBMED 15466573
REFERENCE 2 (bases 1 to 936)
AUTHORS Heu, J.S., Yang, Y.B., Wei, C.L. and Tsai, Y.C.
TITLE Cloning the DAOCS gene from Streptomyces chartreusis
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 936)
REFERENCE Heu, J.S., Yang, Y.B., Wei, C.L. and Tsai, Y.C.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2003) National Yang-Ming University, Institute of
Biochemistry, 115, Li-Nong St. Sec. 2, Shih-Pai, Taipei 112, ROC
FEATURES
source
1..936
/organism="Streptomyces chartreusis"
/mol_type="genomic DNA"
/db_xref="taxon:1969"
1..936
/note="DAOCS"
/codon_start=1
/transl_table=1
/product="deacetoxycephalosporin C synthase"
/protein_id="AAP8381.1"
/db_xref="GI:32492594"
/translation="MDTIVPVFSITELQAGLHQAEPRLCAEGVFLSNSGSEADL
KSKSDVADFFKGTITQKRAVTKADFTIRGFTGLSBSSTAKTNSGSTTDSWCS
MGTSGLNPEHRSAGEWVTGYFDRMYDASRVAQVLKTTETEPGVGVEHVTCEPVR
FRYPPEPERSAEEBPLRMAPHYDLSTVTLIOQTPNPFSVQYVDDGFVDLPNR
PDVLVFCGAVATLATGALVKAAPHNVAAAGROIYVSSRTSVFLRPADFTFSVP
LAKRCGFDIGDGTATFQDMVIMAGNYNLMTKTKA"

ORIGIN
Alignment Scores:
Pred. No.: 1,74e-106 Length: 936
Score: 1300.00 Matches: 238
Percent Similarity: 86.82% Conservative: 32
Best Local Similarity: 76.53% Mismatches: 41
Query Match: 79.46% Indels: 0
DB: 1 Gaps: 0

10719236-1_232-1164 (1-311) x AY318743 (1-936)

QY 1 MetAapThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
DB ATGACACGACGCGTCCGACGCTTACGAGGCTTACGAGGCTCCACACGAGCC 60

QY 21 GluPheArgArgCysLeuArgAspIlysglyLeuPheTyrlleuThraAspCysGlyleuThr 40
DB GAGTCCCGAGCTGCTGCGGAGACAGGTCCTCTTCATCTGACGAGACGCGGCTGTGC 120

QY 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPheGlnHisGlySerGlu 60
DB GAGCCCGCATCAGAAATCCGCGCAAGACGTCGACATTGACTTCTTGCAGACATGCGACG 180

QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
DB GAGGAGAGCGGGCGCGCACCTCCACGATTCCTCCACATTCGCGGGGCTTCACGGGGTTG 240

QY 81 GluSerGluSerThrAlaGlnIleThraPheThrGlySerTyrlleuAspTyrlleuSerMetCys 100
DB GAGTGGAGAGACCGACAGATCACACCGCGGACCTACTCCGACTCTCGATGATGTC 300

QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
DB TACTCGATGGGCTGGCGGACAGATCTTCCCGCGGGGACTTTCAGACGGGCTGTGACG 337

Db 301 TACTCCATGGGAGCACTCCGCAACCTGTTCCCAACGGCGGAGCTTCGAGGCGGTATGAGCG 360
QY 121 GINTYrPhaAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db 361 GGCTACTTCGATCGCATGTACAGAGCGCTCCCGGAGGTGGCCCGGAGGCTCGAAGAGCG 420
QY 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db 421 ACCGGACCGAGCGGAGCGCGGTCTCGAGCACTTGTCTACCTCGCAACCCGCTGCGCG 480
QY 161 PheArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMet 180
Db 481 TTCGGGTACTTCCTCCGAGGTCCTCCGAGCACCGGACGGCGGAGGAGGAGCCCTGCGAGT 540
QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlyThrProCysAlaAsnGly 200
Db 541 GGGCGGCACTACGACTGTGACGCGTACGCTTATCAACAGACCGGTCGCCCAACGGCG 600
QY 201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db 601 TTCGTAGGCTCGAGGTCCAGGTCCAGCGCGGCTTGTGTGACTCGCCGCGCAGCGCGAGC 660
QY 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValIleValAla 240
Db 661 CTGGTCTGCTCTTCTGCGGGGCGGTGCGGACCTTCGCGACCGGCGGCGGTGGTGAAGCA 720
QY 241 ProArgHisIleValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 721 CCACGGCACCGATCCCGCACCGGCGCGGATGATGATGCGGCGAGCGCGTATCGTCC 780
QY 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgIle 280
Db 781 AGCGTCTTCTCTCGGCTCCCAACCGGACCTTCACTTCGCTCCCTCGGCGCAAGCGCG 840
QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrrIleGlyGly 300
Db 841 TCGGATTCGACATCGGCTCGAGCGGCGACAGCGGACAGCTTCAGAGACTCGATCGGAGC 900
QY 301 AenTyrValAsnIleArgArgThrSerIleValAla 311
Db 901 AACTACGTCAACCTCGGTACGAGAACCAAGGCG 933

RESULT 8
A42999 939 bp DNA linear PAT 06-MAR-1997
LOCUS DEFINITION Sequence 14 from Patent WO9504149.
ACCESSION A42999
VERSION A42999.1 GI:2298443
KEYWORDS
SOURCE
ORGANISM
Amycolatopsis lactamdurans
Amycolatopsis lactamdurans
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
REFERENCE
AUTHORS Bovenberg, R.A., Koekman, B.P., Hoekema, A., Van, D.L. and Verweij, J.
TITLE 1 (bases 1 to 939)
JOURNAL PROCESS FOR THE EFFICIENT PRODUCTION OF 7-ADCA VIA
3-(CARBOXYETHYLTHIO)PROPIONYL-7-ADCA
PATENT: WO 9504149-A 14 09-FEB-1995;
GIST BROCADES NV (NL)
Other publication PL 312747 960513
COMMENT Other publication CA 2168004 950209.
FEATURES
location/Qualifiers
1..939
/organism="Amycolatopsis lactamdurans"
/mol_type="unassigned DNA"
/isolate="ATCC 27382"
/db_xref="taxon:1913"
1..939
/gene="CFE"
1..>939
/gene="CFE"
/codon_start=1
/transl_table=11

/protein_id="CAA02697.1"
/db_xref="GI:2298444"
/translation="MTDAITVFPDLAEIRBGLHGEFFHNCIREKGVFLTKGTGLAAD
HASAEIADVDFDGTAEKKAWMTPITIRGVALGSESESTAQITNKGKTDYSMSY
SMGTADNLFPSAEPEKAWEDYFARMYRASQVAVLGVSGAEPEVGDALFDCBPL
RLRYPEVEDRVAEEQPLRMAPHYDLITVLHPTPCANGFVSLQVDFGSDYVDI PA
OPGALVCGAVATVADGAIKAPRHNHVAAGADKRVSSRSVSFPLRPNDGFRFSV
PRARCGDVSIPIAETATFDDWIGGNINIKTAAR"

ORIGIN

Alignment Scores:

Pred. No.:	1,84e-98	Length:	939
Score:	1210.00	Matches:	220
Percent Similarity:	83.87%	Conservative:	40
Best Local Similarity:	70.97%	Mismatches:	50
Query Match:	73.96%	Indels:	0
DB:	6	Gaps:	0

10719236-1_232-1164 (1-311) x A42999 (1-939)

QY 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGlu 21
Db 7 GACGGACCGGTGCGACCTTCGATCTGCGCGAGTGCATGAGGGCTTCACACGAGAGAG 66
QY 22 PheArgArgCysLeuArgAspIleAspIleuPheTyrIleuThrAspCysGlyLeuThrAsp 41
Db 67 TTCGCGACCTGCTCGGAGAGAGGGCGGTGTCTACTCTCAAGGAGCAACGGGCTCCCGAG 126
QY 42 ThrGluLeuYsSerAlaIleAspIleValIleAspPhePheGluHisGlySerGluAla 61
Db 127 GCGGACCAACGCTGCGCGGAGAGATCGCGGTGACCTTCGACCAAGGACCGAGGCG 186
QY 62 GluIleValAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlu 81
Db 187 GAGAAAGAGCGCGGTGATACCGCATCCGACCATCCGCGCGGCTGACCGCGGCTGAG 246
QY 82 SerGluSerThrAlaGlnIleThrAspThrGlySerTyrSerAspTyrSerMetCysTyr 101
Db 247 TCGGAGACACCGGCAATCAAGACACCGCAAGTACACGACTTACTCGATGCTGTC 306
QY 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTrrPThrGln 121
Db 307 TCGATGGGACCGCGGACACACTGTTCCACGCGCGGACGTTGAGAAAGCGGTGGAGGAC 366
QY 122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
Db 367 TACTTCGCGGATGTACCGCGCTTCGCGACGTCGCGCGGAGGTGCTGACCTCGGTC 426
QY 142 GlyThrGluProAspGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
Db 427 GCGCGGAAACCGAGGTGCGGATGGAACGCTTCTCGACTGCGAAACCTCGTGTGCGCTG 486
QY 162 ArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAla 181
Db 487 CGCTACTTCCCGAGAGTCCCGAGGATGCGGTGCGGAGGAGGACCGCTGCGATGCGC 546
QY 487 ProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPhe 201
Db 547 CCGGACTACGACCTTCGATCGTACACCTGTATCCACCGAACCCCTTGTGCGCAACCGGTTTC 606
QY 202 ValSerLeuGlnAlaGluValGlyValaPheThrAspLeuProTyrArgProAspAla 221
Db 607 GTGACCTGACGCTGAGGTGAGACGCGTCTATGTGACATCCGCGCGAGCGGCGCG 666
QY 222 ValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValIleValPro 241
Db 667 GTGCTGCTGCTCTGCGGCGGCGGTGCGACGCTGTGTGCGGACGCGCGCATCAAGCGCGC 726
QY 242 ArgHisIleValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
Db 727 AAGCACCACTGTCGCGCGCGCGGCGCGGATGATGATGATGATGATGATGATGATGATG 786
QY 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCys 281

Db 787 GTGTTCTTCTGCGCCCAAGCGGAGCTTCGCTTCTGCTGCGCGGAGCCAGGAGTGC 846

Qy 282 GlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyValAsn 301

Db 847 GGGTTCGACGTCAGATCCCGGCGGAGCCGACCTTCGACGACTGATCGGCGGCAAC 906

Qy 302 TyrValAsnIleArgArgThrSerIysAla 311

Db 907 TACATCAACATCCGGAAGACCGCGCGCGC 936

RESULT 9

LOCUS A72870 939 bp DNA linear PAT 15-OCT-1999

DEFINITION Sequence 14 from Patent WO9504148.

ACCESSION A72870

VERSION A72870.1 GI:6063888

KEYWORDS

SOURCE Amycolatopsis lactandurans

ORGANISM Amycolatopsis lactandurans

REFERENCE Bacteria; Actinobacterii; Actinobacteridae; Actinomycetales; Pseudonocardineae; Pseudonocardiaceae; Amycolatopsis.

AUTHORS Bovenberg, R.A. and Koekman, B.P.

TITLE PROCESS FOR THE EFFICIENT PRODUCTION OF 7-ADCA VIA 2-(CARBOXYETHYLTHIO)ACETYL-7-ADCA AND 3-(CARBOXYMETHYLTHIO) PROPIONYL-7-ADCA

JOURNAL Patent: WO 9504148-A 14 09-FEB-1995; Location/Qualifiers

FEATURES

source 1..939

1..939

/organism="Amycolatopsis lactandurans"

/mol_type="unassigned DNA"

/isolate="ATCC 27382"

/db_xref="taxon:1913"

1..939

/gene="CEFE"

1..>939

/gene="CEFE"

/codon_start=1

/transl_table=11

/protein_id="CAB58463.1"

/db_xref="GI:6063888"

/translation="MTDATVPTPDLAELREGHDEPHRCLEKGVFLKGTGABAD HASAREIVDPEDHGTBAKKVMTPTIRGRGAGLESTQITNGKXTDYSMY SMGTADNLPPSAFEKAWEDYFARKYRASQDYARQVLTSGAEPEVGDALDCEPLL RLRYPEVEDEVRVAEOPLRMAPHYDLSIVTLHQTPANGPVLSQVEVDGSYVDIPA OPGAVLFCGAVATLVADGAIKAPKHHVAAGADKRVGSSRTSSVFLRPNGDFRSY PRARCGDVSIAPETATFDWIGNYINIRKTAAR"

ORIGIN

Alignment Scores:

Pred. No.: 1.84e-98 Length: 939

Score: 1210.00 Matches: 220

Percent Similarity: 83.87% Conservative: 40

Best Local Similarity: 70.97% Mismatches: 50

Query Match: 73.96% Indels: 0

DB: 6 Gaps: 0

10719236-1_232-1164 (1-311) x A72870 (1-939)

Qy 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHiIeGlnAspGlu 21

Db 7 GACGGACCGCTGCGACCTTCATCTGCGCGAGCTGCGGAGGCTTGCACCAAGAGAG 66

Qy 22 PheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeuThrAsp 41

Db 67 TTCGCCACATGCTGCGGAGGAGGCGGTCTTCACTCAAGGAGCAGCGGCTCGCCGAG 126

Qy 42 ThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGluHiIeGlySerGluAla 61

Db 127 GCGGACCAAGCTCGCGCGGAGATCGCGTGAATCTTTCGACCAAGGACGAGGCC 186

Qy 62 GluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlu 81

Db 187 GAGAAAGAGCGGTGATGACGCCGATCCGACCATCCGCGCGGGTACGCCGGCTGGAG 246

Qy 82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyr 101

Db 247 TCCGAGAGACCGCGGATCAGCAACACCGGCAAGTACACCGACTCTGATGTCTGATC 306

Qy 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluAlaGlyIleTyrThrGln 121

Db 307 TCGATGGGACCGCGGACCAACTGTTCCGAGCGCGGATTCGAGAAAGCGGTGGAGAC 366

Qy 122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141

Db 367 TACTTCCGCGGAGTACCGCGCTTCGACAGACTCCGCGGAGAGTGTGACTCGGTC 426

Qy 142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161

Db 427 GCGCGGACCGCGGAGTCCGCGATGACGCTTCTCGACTGCGAACCCCTGCTGCGCTG 486

Qy 162 ArgTyrPheProGlnValProGluHiIeArgSerAlaGlnGluGlnProLeuArgMetAla 181

Db 487 CGTACTTCCCGCGAGTCCCGAGAGATCGGTGCGCGAGGAGACCGCTGCGGATGGCC 546

Qy 182 ProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPhe 201

Db 547 CCGGACTACGACCTCTCGATGCTACCTGATCCACGACACCCCTTGCGGAAAGGCTTC 606

Qy 202 ValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAla 221

Db 607 GTCAGCTCGAGGTCGAGGTGAGACGCGGTCTATGTGACATCCCGGCGGAGCGGCGCG 666

Qy 222 ValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGluValAlaPro 241

Db 667 GTGCTGGTGTCTGCGCGCGGTGGCAAGCTGTGTGCCACAGCGCGGATCAAGGCCCC 726

Qy 242 ArgHisIleValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261

Db 727 AAGCACCAAGCGCGCGCGCGCGCGCGGAGCAAGCGGAGTGGGACGACCCGACCTCCAGC 786

Qy 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCys 281

Db 787 GTGTTCTTCTGCGCCCAAGCGGAGCTTCGCTTCTGCGTGGCGCGGCGGAGGAGTGC 846

Qy 282 GlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyValAsn 301

Db 847 GGGTTCGACGTCAGATCCCGGCGGAGCCGACCTTCGACGACTGATCGGCGGCAAC 906

Qy 302 TyrValAsnIleArgArgThrSerIysAla 311

Db 907 TACATCAACATCCGGAAGACCGCGCGCGC 936

RESULT 10

LOCUS AR023763 939 bp DNA linear PAT 05-DEC-1998

DEFINITION Sequence 15 from patent US 5795733.

ACCESSION AR023763

VERSION AR023763.1 GI:3977057

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

AUTHORS Bovenberg, R.Ary Lams., Koekman, B.Pieter., Hoekema, A., Van Der Laan, J.Meiske., Verweij, J. and De Vroom, E.

TITLE Process for the efficient production of 7-ADCA via 3-(carboxyethylthio) propionyl-7-ADCA

JOURNAL Patent: US 5795733-A 15 18-AUG-1998; Location/Qualifiers

FEATURES

source 1..939

1..939

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.:	1,84e-98	Length:	939
Score:	1210.00	Matches:	220
Percent Similarity:	83.87%	Conservative:	40
Best Local Similarity:	70.97%	Mismatches:	50
Query Match:	73.96%	Indels:	0
DB:	6	Gaps:	0

10719236-1_232-1164 (1-311) x AR023763 (1-939)

```
QY 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGlu 21
   7 GACGGACCGCTGCCGACCTTCGATCTGCCGAGCTCGTAGGGCTTCGACCGAGAGAG 66
QY 22 PheArgArgCysLeuArgAspGlyGlyLeuPheTyrLeuThrAspCysGlyLeuThrAsp 41
   67 TTCCTCCGACCTGCTCGCGCAGAGAGGCGTCTTCTACCTCAAGGGCACCGGGCTCGCCGAG 126
QY 42 ThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySerGluAla 61
   127 GCGGACACCGCTCGCGCGGAGATCGCGTGACTTCTTCGACCAACGCGACCGAGGCC 186
QY 62 GluLysArgAlaValThrSerProValProThrMetArgGlyPheThrGlyLeuGlu 81
   187 GAGAGACGCGGTGATGACCGCATCCGACCATCCGCGCGGTACCGCGGCTCGAG 246
QY 82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyr 101
   247 TCCGAGACGACCGCGCAGATCAACGAGACCGGCAAGTACACCACTACCTGATGTCGTAC 306
QY 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThrGln 121
   307 TCGATGGGACCGCGGACAACTGTCTCCAGCGCGGAGTTCAGAAAGCGGTGGAGAGAC 366
QY 122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
   367 TACTTGGCGGAGATGATACCGCGCTTCGAGAGAGTGGCGGGAGGTGTCGACTCGGTC 426
QY 142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
   427 GCGCGGAGAACCGGATCGGATGAGACGCGCTTCTGACTGCGAAACCTCGTGGCCCTG 486
QY 162 ArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAla 181
   487 CGCTACTTCCCGAGGTGCGCGAGATCGCGTGGCGGAGAGACGCGTGGATGGGC 546
QY 182 ProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPhe 201
   547 CCGGACTACGACCTTCGATCGTACACCTGATCCACAGACCCCTTGGCGGAGCGGATTC 606
QY 202 ValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAla 221
   607 GTACAGCTCGAGGTGAGAGGTGAGCGGGTCTTGTGATCATCCGGCGGACCGGGCGGG 666
QY 222 ValLeuValPheCysGlyValAlaIleAsnThrLeuValThrGlyGlyGlnValAspAlaPro 241
   667 GTGCTGTGTCTTCGCGCGCGGTGGTGGAGCGTGTGGCCGAGCGCGGTCAAGGGCGCC 726
QY 242 ArgHisHisValAlaAlaIleProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
   727 AAGGACCAAGTGGCGCGCGCGCGCGGAGCAACGCGGTGGGAGGAGCGCGACCTTCAGC 786
QY 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlyCys 281
   787 GTGTCTTCTCTCGCGCGCGGAGCTTCGCTTCTCGGTGCGCGGCGCGAGGGAGTGC 846
QY 282 GlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGlyAsn 301
   847 GGGTTCGAGCTGACGATCCCGCGCGGAGACCGGACCTTCGACGACTGATCGCGGAGAC 906
QY 302 TyrValAsnIleArgArgThrSerIleValAla 311
   |||
```

```
DB 907 TACATCAATCCGGAAGACCGCGCGGCC 936
RESULT 11
LOCUS 191794 939 bp DNA linear PAT 01-DEC-1998
DEFINITION Sequence 14 from patent US 5726032.
ACCESSION 191794
VERSION 191794.1 GI:3936264
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 939)
AUTHORS Bovenberg, R.Ary, Ians, Koekman, B. Pieter, Hoekema, A., Van Der
Laan, J., Mecke, Verweij, J., and De Vroom, E.
TITLE Process for the efficient production of 7-ADCA via
2-(carboxymethylthio)acetyl-7-ADCA and
3-(carboxymethylthio)propionyl-7-ADCA
JOURNAL Patent: US 5726032-A 14 10-MAR-1998;
FEATURES
source location/Qualifiers
1..939
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,84e-98 Length: 939
Score: 1210.00 Matches: 220
Percent Similarity: 83.87% Conservative: 40
Best Local Similarity: 70.97% Mismatches: 50
Query Match: 73.96% Indels: 0
DB: 6 Gaps: 0
10719236-1_232-1164 (1-311) x 191794 (1-939)
QY 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGlu 21
   7 GACGGACCGCTGCCGACCTTCGATCTGCCGAGCTCGTAGGGCTTCGACCGAGAGAG 66
QY 22 PheArgArgCysLeuArgAspGlyGlyLeuPheTyrLeuThrAspCysGlyLeuThrAsp 41
   67 TTCCTCCGACCTGCTCGCGCAGAGAGGCGTCTTCTACCTCAAGGGCACCGGGCTCGCCGAG 126
QY 42 ThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySerGluAla 61
   127 GCGGACACCGCTCGCGCGGAGATCGCGTGACTTCTTCGACCAACGCGACCGAGGCC 186
QY 62 GluLysArgAlaValThrSerProValProThrMetArgGlyPheThrGlyLeuGlu 81
   187 GAGAGACGCGGTGATGACCGCATCCGACCATCCGCGCGGTACCGCGGCTCGAG 246
QY 82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyr 101
   247 TCCGAGACGACCGCGCAGATCAACGAGACCGGCAAGTACACCACTACCTGATGTCGTAC 306
QY 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThrGln 121
   307 TCGATGGGACCGCGGACAACTGTCTCCAGCGCGGAGTTCAGAAAGCGGTGGAGAGAC 366
QY 122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
   367 TACTTGGCGGAGATGATACCGCGCTTCGAGAGAGTGGCGGGAGGTGTCGACTCGGTC 426
QY 142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
   427 GCGCGGAGAACCGGATCGGATGAGACGCGCTTCTGACTGCGAAACCTCGTGGCCCTG 486
QY 162 ArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAla 181
   487 CGCTACTTCCCGAGGTGCGCGAGATCGCGTGGCGGAGAGACGCGTGGATGGGC 546
QY 182 ProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPhe 201
   |||
```

Db 547 CCGACCTACGACCTTCGATCGTCACCTCGATCCACGACCCCTTGCGGACGGGTTC 606
Qy 202 ValSerLeuGlnAlaGluValGlyAlaPheThrAspLeuProTyrArgProAspAla 221
Db 607 GTCAAGCTCGACAGGTGAGGTGGACGGGTCTATGTGACATCCCGGACACCGGCGCG 666
Qy 222 ValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValValAlaPro 241
Db 667 GTGCTGTTCTCTGGCGCGGTGGCGACGGCTGGCGCCACGGCGGCATCAAGCGGCC 726
Qy 242 ArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
Db 727 AAGCACCAAGTGGCGCGCGCGCGGACCAAGCGGGTGGGACGACCGACCTCCACG 786
Qy 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlyCys 281
Db 787 GTGTTCTTCTCTGGCGCGCGGACGCTTCCGCTTCTGGTCCCGGCGGACGAGAGTC 846
Qy 282 GlyPheAspValSerLeuAspGlyGlnThrAlaThrPheGlnAspTyrIleGlyGlyAsn 301
Db 847 GGGTTCGACGTACGATCCCGCGCGGACCGCACCTTTCGACGACTGGATCGCGGCAAC 906
Qy 302 TyrValAsnIleArgArgThrSerTyrAla 311
Db 907 TACATCAACATCCGGAAGACCGCGCGCGC 936
RESULT 12
A43001 A43001 942 bp DNA linear PAT 06-MAR-1997
LOCUS Sequence 16 from Patent WO9504149.
DEFINITION A43001
ACCESSION A43001.1 GI:2298445
KEYWORDS
ORGANISM Amycolatopsis lactamdurans
SOURCE Amycolatopsis lactamdurans
Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
Pseudonocardineae; Pseudonocardiaceae; Amycolatopsis.
REFERENCE 1 (bases 1 to 942)
Bovenberg, R.A., Koekman, B.P., Hoekema, A., Van, D.L. and Verweij, J.
PROCESS FOR THE EFFICIENT PRODUCTION OF 7-ADCA VIA
3-(CARBOXYETHYLTHIO)PROPIONYL-7-ADCA
Patent: WO 9504149-A 16 09-FEB-1995;
GIST BROCADES NV (NL)
COMMENT Other publication PL 312747 960513
Location/Qualifiers
FEATURES
source
1..942
/organism="Amycolatopsis lactamdurans"
/mol_type="unassigned DNA"
/isolate="IC 411"
/db_xref="taxon:1913"
1..942
/gene="CFE"
/gene="CFE"
1..>942
/gene="CFE"
/codon_start=1
/transl_table=11
/protein_id="CAA02698.1"
/db_xref="GI:2298446"
/translation="MTDAVPTFPDLAEIRREGIHOEFPHCLREKGVFLKGTGPAEA
DHASGREIADVPFDGTEAEKKAVMPTPIPTIRGVAGSESTAOITWTGKVTDSMS
YSMTGADNLPSPAEPEKAMEDYFARMYASODVAROVLTSVGAPEVGMADFLDEPL
LRLRFPFPEPRVAEBOPLRMAHYDLSTVTLHQTCANGFVSLQYEVDSYVDIP
AOPGAVLFCGAVTLVADGAIKAPKHVAAPGADKRVGSSRTSSVFLLRPVDFRFS
VPRARECGPDAVIPAETATFDWIGNYINIRKTAAR"

DB: 6 Gaps: 1
10719236-1_232-1164 (1-311) x A43001 (1-942)
Qy 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGly 21
Db 7 GAGCGACCGGTGGACCTTCGATCTGGCGAGCTGGTGAAGGCTTGACACGAGAGAG 66
Qy 22 PheArgArgCysLeuAlaAspGlyGlyLeuPheTyrLeuThrAspCysGlyLeu---Thr 40
Db 67 TTCGCGACATGCTCGCGCGGAGAGGGGTGTCTTACTCAAGGACCGGAGCTCGGCC 126
Qy 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGlnHisGlySerGlu 60
Db 127 GAGCGACCAACGCTCGCGCGCGGAGATCGCGGTGATCTTTCACACACGACCGAG 166
Qy 61 AlaGluLysArgAlaValThrSerProValProThrMetArgGlyPheThrGlyLeu 80
Db 187 GCCGAGAAAGAGGGGTGATGACCGCGATCCGACATCCGCGCGGTACGCGCGGCTG 246
Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db 187 GCCGAGAAAGAGGGGTGATGACCGCGATCCGACATCCGCGCGGTACGCGCGGCTG 246
Qy 247 GAGTCCAGAGACCGCGCGGACATCAGAACCGGCAAGTACACCGACTCTCATTCG 306
Db 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
Qy 307 TACTCGATGGGACCGCGCGGACCACTTTCACGCGCGGACAGTTTCAGAAAGCGTGGAG 366
Db 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Qy 367 GACTACTTCGCGCGGAGTACCGCGCTTCGACAGACGTCGCGCGGAGTGCATCTCG 426
Db 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Qy 427 GTCCGCGGAAACCGCGGTCGATGAGCGATGACGCTTCTTCATCGACCGAACCTCTCG 486
Db 487 CTGGCTACTCTCCCGAGAGTCCCGAGAGTCCGCTGGCCGACGAGACCGCGCTCGGATG 546
Qy 161 PheArgTyrPheProGlnValProGluHisArgSerAlaGluGlnGlnProLeuArgMet 180
Db 487 CTGGCTACTCTCCCGAGAGTCCCGAGAGTCCGCTGGCCGACGAGACCGCGCTCGGATG 546
Qy 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db 547 GCCCGCACTACGACCTCTCGATGTCACCTGATCCACAGACCCCTTCGCGCAAGCGG 606
Qy 201 PheValSerLeuGlnAlaGluValGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db 607 TTCGTCAGCTGCGAGGTGAGGTGAGCGGCTCTATGTGACATCCCGCGCGCGCGG 666
Qy 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValAla 240
Db 667 GCGGTGCTGTGTCTCTCGCGGCGGTGGCGACGCTGTGGCGGACCGCGGATCAAGGCG 726
Qy 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 727 CCCAAGCACCAAGTGGCGCGCGCGCGGACCAAGCGGTGGCGACAGCGCACGCTCC 786
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db 787 AGCGTGTCTTCTCGCGCGCGCAAGCGGACTTCGCTTCTCGGTGGCGGCGGACGAG 846
Qy 281 CysGlyPheAspValSerLeuAspGlyGlnThrAlaThrPheGlnAspTyrIleGlyGly 300
Db 847 TCGCGGTTCGACGTCGATCCGCGCGGAGACCGGCACTTTCACGACGATCGGCGGC 906
Qy 301 AsnTyrValAsnIleArgArgThrSerTyrAla 311
Db 907 AACTACATCAACATCCGGAAGACCGCGCGCGC 936
RESULT 13
A72872 A72872 942 bp DNA linear PAT 15-OCT-1999
LOCUS Sequence 16 from Patent WO9504148.
DEFINITION A72872
ACCESSION A72872

ALIGNMENT Scores:

Pred. No.: 3,61e-97 Length: 942
Score: 1195.50 Matches: 219
Percent Similarity: 83.28% Conservative: 40
Best Local Similarity: 70.42% Mismatches: 51
Query Match: 73.07% Indels: 1

VERSION A72872.1 GI:6063890
 KEYWORDS
 SOURCE Amycolatopsis lactamdurans
 ORGANISM Amycolatopsis lactamdurans
 Bacteria; Actinobacteria; Actinomycetales;
 Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
 REFERENCE 1 (bases 1 to 942)
 Bovenberg, R.A. and Koekman, B.P.
 TITLE PROCESS FOR THE EFFICIENT PRODUCTION OF 7-ADCA VIA
 2-(CARBOXYETHYLTHIO)ACETYL-7-ADCA AND
 3-(CARBOXYETHYLTHIO)PROPIONYL-7-ADCA
 JOURNAL Patent: WO 9504148-A 16 09-FEB-1995;
 GIST BROCADES NV (NL); BOVENBERG ROELOF ARY IANS (NL)
 FEATURES
 source
 location/Qualifiers
 1..942
 /organism="Amycolatopsis lactamdurans"
 /mol_type="unassigned DNA"
 /isolate="LC 411"
 /db_xref="taxon:1913"
 1..942
 /gene="CERE"
 1..942
 /gene="CERE"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB58464.1"
 /db_xref="GI:6063891"
 /translation="MTDATVPFTDLAELEBGLHQBEPHCLREKGVFLKGLGLPAEA
 DHASGREIADVDFDHTGEAKVMTPIPIIRGVAGLESSTQAQITNGKTYDMS
 YSMGADNLFPSAEFEKAWEDYFARMYASODVARQVLSVSGEPVGMADFLDCPL
 LRLRVPEVDEDEVAEOPFLMAYPHYDLISVTLIHPGANGVSLQVSDSYDIP
 ADGAVLVPFGAVALTVADGAIKAPKHVAALPGADGRVSSSRISVFLRPNDPFS
 VPRARCGDVSIPAEATATPDMTIGNTYINIKTAAR"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,616-97 Length: 942
 Score: 1195.50 Matches: 219
 Percent Similarity: 83.28% Conservative: 40
 Best Local Similarity: 70.42% Mismatches: 51
 Query Match: 73.07% Indels: 1
 DB: 6 Gaps: 1
 10719236-1_232-1164 (1-311) x A72872 (1-942)
 QY 2 AepThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAapGlu 21
 DB 7 GACGCGACCGTGCACCTTCGATCTGGCCGACCTCGTGAGGCGTTGCACGAGAGAG 66
 QY 22 PheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeu--Thr 40
 DB 67 TTCCGCGACCTGCTGGCGAAGGCGGTCTTACCTCAAGGGCACCGGCTGCCGCC 126
 QY 41 AepThrGluLeuSerAlaValAlaThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
 DB 127 GAGCGCGACACCGCTTCGGCCGCGAGATCGCGTGAATCTTCCACACGCGACCGAG 186
 QY 61 AlaGluLeuArgAlaValAlaThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
 DB 187 GCCGGAAGAAGCGCGGTGATGACGCGGATCCGACCATCCGCGCGGATACCGCGGGCTG 246
 QY 81 GluSerGluSerThrAlaGlnIleThrThrThrGlySerTyrSerArgTyrSerMetCys 100
 DB 247 GAGTCCGAGAGCACCCGCGAGATCACGAACACCGGCAAGTACCCGACTACTCGAGTGTG 306
 QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluIuArgIleTPTPr 120
 DB 307 TACTGATGGGCAACCGCGGACACTGTTCCGAGCGCGAGTTCGAGAAAGCGGTGGAG 366
 QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
 DB 367 GACTACTTCGCGCGAGATGATACCGCGCTTCGCGAGAGAGTGTGACTGACCTCG 426

QY 141 ThrGlyThrGluProAspGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
 DB 427 GTCCGCGCGGAACCCGAGGTCCGATGACGCTTCTTCGACTGCGAACCCTGCTGCC 486
 QY 161 PheArgTyrPheProGluAlaProGluHisArgSerAlaGluGluGlnProLeuArgMet 180
 DB 487 CTGCGCTACTTCCCGAGAGTCCCGAGATGCGTGGCCGAGACGACGCGCTCGGATG 546
 QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
 DB 547 GCCCGGACATCAAGACCTTCGATCGTCAACCCGATCCACGACCCCTTGGCCGAAACGG 606
 QY 201 PheValSerLeuGluAlaGluValAlaGlyValAlaPheThrAspLeuProTyrArgProAsp 220
 DB 607 TTTCGACGCTTCGACGCTCGAGGTGACGCGGTCTTATGTCGACATCCCGCGACCGGCG 666
 QY 221 AlaValLeuValPheCysGlyValAlaIleAlaThrLeuValThrGlyGlyValIysAla 240
 DB 667 GCGGTGCTGTGTCTTCGCGCGCGGTGGAGCGTGTGGCCGAGCGCGCATCAAGCGCG 726
 QY 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
 DB 727 CCCAAGCACACAGTGGCGCGCGCGCGCGGACAAAGCGGTGGGACGACGCGCACTTC 786
 QY 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
 DB 787 ACGGATTTCTTCTGCGCCCAACGGGACCTTCGCTTCGATGCGCGCGCGCGCGAG 846
 QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300
 DB 847 TCGGAGTTTCGACGTGACGATCCCGGCGGACGCGCACCTTCGACGACGTGATCGCGCG 906
 QY 301 AenTyrValAsnIleArgArgThrSerTyrAla 311
 DB 907 AACTACATCAACATCCGGAAGACCGCGCGCGCG 939
 RESULT 14
 LOCUS AR023762 942 bp DNA linear PAT 05-DEC-1998
 DEFINITION Sequence 14 from patent US 5795733.
 ACCESSION AR023762
 VERSION AR023762.1 GI:3977056
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 942)
 Bovenberg, R.Ary.Ians., Koekman, B.Pieter., Hoekema, A., Van Der
 Laan, J.Metske., Verweij, J. and De Vroom, E.
 TITLE Process for the efficient production of 7-ADCA via
 3-(carboxyethylthio) propionyl-7-ADCA
 JOURNAL Patent: US 5795733-A 14 18-AUG-1998;
 FEATUES location/Qualifiers
 source 1..942
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,616-97 Length: 942
 Score: 1195.50 Matches: 219
 Percent Similarity: 83.28% Conservative: 40
 Best Local Similarity: 70.42% Mismatches: 51
 Query Match: 73.07% Indels: 1
 DB: 6 Gaps: 1
 10719236-1_232-1164 (1-311) x AR023762 (1-942)
 QY 2 AepThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAapGlu 21
 DB 7 GACGCGACCGTGCACCTTCGATCTGGCCGACCTCGTGAGGCGTTGCACGAGAGAG 66
 QY 22 PheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeu--Thr 40

```
Db 67 TTCCCACTGCTCTCGGAGAGAGCGGCTTCTCACTTCAAGGCAACCGGCTGCCGCC 126
Qy 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySerGlu 60
Db 127 GAGGCGACACACGCTCGGCGCGGAGATCGCGGTGACTTCTTGCACCGCACCGAG 186
Qy 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db 187 GCCGAGAAAGAGCGCGTGTATGACCGCATCCGACCATCCGCGCGGGTATCGCCGGGCTG 246
Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db 247 GAGTCCGAGAGACCGCGCGAGATCAAGAACCGGCAAGTACACCGACTACTCGATGTCG 306
Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTrpThr 120
Db 307 TACTCGATGGGACCGCGGACAACCTGTCCCAAGCGCGAGTTCGAGAAGCGTGGAG 366
Qy 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db 367 GACTACTTCGCGCGGATGTACCGCGCTTCGACGACGTGCGCGGAGGTGTGACCTCG 426
Qy 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGlyProLeuLeuArg 160
Db 427 CTCGCGCGCGAACCAGAGTGGCATGACGCTTCTCGACTGCGAACCCTGCTGCGC 486
Qy 161 PheArgTyrPheProGlnValProGlnHisArgSerAlaGluGlnProLeuArgMet 180
Db 487 CTGCGCTACTTCCCGAGGTGCGCCAGAGATCCGCGCGAGACAGCCGCTGCGAGT 546
Qy 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db 547 GCCCGGACATGACCTCTCGATGCGACCTCGATCACAGAACCCCTTGCGCGAGACG 606
Qy 201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db 607 TTCGTGAGCTGTCCAGGTGAGGTGACGCGCTCTATGTGACATCCCGCGCGAGCGG 666
Qy 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValIlyAla 240
Db 667 GCGGTGCTGTGTCTGCGCGCGGTGTGCGACGTGTGTGCGCGCGCGGTCAAGGCG 726
Qy 241 ProArgHisHisValAlaAlaProArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 727 CCCAAGACACACGTGCGCGCGCGCGCGAGACGCGGTGGAGGACGCGCACCTCC 786
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db 787 ACGGTGTTCTTCTGTGCGCCCAACGGGACCTTCCGTTCTGTGTCGCGCGGCGCAG 846
Qy 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300
Db 847 TCCGGGTTCGACGTGCGCATCCCGCGCGAGCCGCACTTCGACGACTGATCGCGCG 906
Qy 301 AsnTyrValAsnIleArgArgThrSerLysAla 311
Db 907 AACTACATCAACATCCGAGAACCGCGCGGCC 939

RESULT 15
LOCUS 191795 942 bp DNA linear PAT 01-DEC-1998
DEFINITION Sequence 16 from patent US 5726032.
ACCESSION 191795
VERSION 191795.1 GI:3936265
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 942)
AUTHORS Bovenberg, R.Ary.Jans., Koekman, B.Pieter., Hoekema, A., Van Der
Laan, J.Melke., Verweij, J., and De Vroom, E.
TITLE Process for the efficient production of 7-ADCA via
```

```
2-(carboxyethylthio)acetyl-7-ADCA and
3-(carboxymethylthio)propionyl-7-ADCA
Patent: US 5726032-A 16 10-MAR-1998;
FEATURES
source
1. .942
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3,61e-97 Length: 942
Score: 1195.50 Matches: 219
Percent Similarity: 83.28% Conservative: 40
Best Local Similarity: 70.42% Mismatches: 51
Query Match: 73.07% Indels: 1
Gaps: 1
10719236-1_232-1164 (1-311) x 191795 (1-942)
Qy 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnIleuHisGlnAspGlu 21
Db 7 GACCGACCGTGGCCGACCTTCGATCTGACGCGAGCTGTGAAGGCTTGACACGAGAG 66
Qy 22 PheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeu---Thr 40
Db 67 TTCCGCACTCCCTCGCGCGAGAAAGGCGTGTCTTACTCTCAAGGSCAACCGGCTGCCGCC 126
Qy 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySerGlu 60
Db 127 GAGCGGACACACGCGCTGGCGCGGAGATCGCGGTGACATCTTTCGACACCGGACGAG 186
Qy 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db 187 GCCGAGAAAGAGCGGTGATGACCCCGATCCGACCATCCGCGCGGCTTACCGCGGCTG 246
Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db 247 GAGTCCGAGAGACCGCGGACGATCAGAACACCGGCAAGTACACCGACTACTCGATGTCG 306
Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTrpThr 120
Db 307 TACTCGATGGGACCGCGGACCAACCTTCTCCCAAGCGCGGTTCGAGAAAGCGTGGAG 366
Qy 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db 367 GACTACTTCGCGCGGAGTACCGGCTTTCGACGACGTCCGCGCGAGTGTGACTCGC 426
Qy 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGlyProLeuLeuArg 160
Db 427 GTCGCGCGGAAACCGAGGTGCGCATGACGCTTCTTCGACTGCGAACCCTGCTGCGC 486
Qy 161 PheArgTyrPheProGlnValProGlnHisArgSerAlaGluGlnProLeuArgMet 180
Db 487 CTGCGCTACTTCCCGAGGTGCGCCAGAGATCGCTGTGCGAGGACGCGGTGGATG 546
Qy 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db 547 GCCCGGACATGACCTCTCGATGTCACCTTATCAACGACACCTTTCGCGCAACGGG 606
Qy 201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db 607 TTCGTGAGCTGTGCGAGGTGAGGTGAGCGGCTTATGTGACATCCCGCGCGACCGGCG 666
Qy 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValIlyAla 240
Db 667 GCGGTGCTGTGTCTCGCGCGGTGTGCGACGCTGTGTGCGCGCGCGGTCAAGGCG 726
Qy 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 727 CCCAAGACACACGTGCGCGCGCGCGCGGACGAGGCGGTGGACACCGCACTCC 786
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
```

```

Db      787 AGCGTGTCTTCTCGGCCCAACGGGGAATTCCGCTCTCGGTGCCCGGCCAGGAG 846
QY      281 CveGlyPheAspValSerLeuAseGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300
Db      847 TCGGGGTTCGACGTGACATCCCGGCCGAGACCGCACCTTCGACGACTGGATCGGGC 906
QY      301 AsnTyValAsnIleArgArgThrSerTyAla 311
Db      907 AACTACATCAACATCCGGAAGACCGCCGCCGCC 939

```

Search completed: September 19, 2005, 16:57:53
 Job time : 4929 secs

THIS PAGE BLANK (USPTO)


```

XX  Yang Y, Wei C, Hsu J, Tsai Y;
XX  WPI; 2003-78116/74.
XX  P-PSDB; ADE47638.
XX
XX  New mutated penicillin expandase having an expandase activity towards
XX  penicillin G, which is at least at 2-fold higher than the activity of
XX  wild-type expandase, useful for producing 7-aminodesacetoxycephalosporanic
XX  acid (7-ADCA).
XX
XX  Disclosure; SEQ ID NO 1; 55bp; English.
XX
XX  The invention relates to a novel mutated penicillin expandase having an
XX  expandase activity towards penicillin G, which is at least at 2-fold
XX  higher than the activity of wild-type expandase. The mutated penicillin
XX  expandase is useful for producing 7-aminodesacetoxycephalosporanic acid.
XX  The present sequence encodes the wild-type penicillin expandase.
XX
XX  Sequence 1230 BP; 207 A; 455 C; 384 G; 184 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      6,436-167      Length:      1230
Score:          1636.00      Matches:      311
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             10      Gaps:         0

10719236-1_232-1164 (1-311) x ADE47637 (1-1230)

QY  1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAsp 20
DB  232 ATGGACAGACGGCTGCCCACTTCAGCCCTGGCCGAACTCCAGCAGGCGCTGACCAAGAC 291
QY  21 GluPheAArgCysGlyLeuAArgAspLyseGlyLeuPheTyLeuThrAspCysGlyLeuThr 40
DB  292 GAGTTCGCCAGGTGCTCTGAGGGACAAGGGCCCTCTTCTATCTGACGCACTGGCGGTGTGACC 351
QY  41 AspThrGluLeuLyseSerAlaLysAspIleValIleAspPhePheGluHisGlySerGlu 60
DB  352 GACACCGAGCTGAACTCGCGCCAAAGACATCTGTCATCGACTTCTTTCGAGCAGCGCGCCAG 411
QY  61 AlaGluLyseArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
DB  412 GCGGAAAGCGCGCGCTGACCTCGCCGCCGCCACCAATGCGCGGCGCTTCCACCGAGCTG 471
QY  81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTySerAspTySerMetCys 100
DB  472 GAGTCGAGAGACACCGCCCAAGATCAACAATACCGGCAAGTACTCCGACTCATCGATGAC 531
QY  101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleThrThr 120
DB  532 TACTCGATGGGACACCGCGCAACCTCTTCCTCGGCTCGGACTTGGAGCGGATCTGAGACC 591
QY  121 GlnTyrrPheAspArgGlnTyrrThrAlaSerArgAlaValAlaArgGluValIleuArgAla 140
DB  592 CAGTACTTCGACCGCCAGTACACCGCTCTCCGCGGGTGTGCCGGAAGGTCTTGGCGGCG 651
QY  141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
DB  652 ACCGGGACCGGACCGCGAGCGGGGTCCAGGCTTCTCTGACTGGAGACCGCTGTGGCG 711
QY  161 PheArgTyrrPheProGlnValProGluHisArgSerAlaGluGlnGlnProLeuArgMet 180
DB  712 TTCCCTCTCTCTCCCGCAGGTCTCCCGAGGACCCGACGCGCGAGAGACGCTTGGGATG 771
QY  181 AlaProHisTyrrAspLeuSerMetValThrIleuIleGlnGlnThrProCysAlaAsnGly 200
DB  772 GCGCGGCACTACGACTGTGATGATGATCCCTCATCTCAGCAACACCTTGGCCCAAGGC 831
QY  201 PheValSerLeuGlnAlaGluValGlyAlaPheThrAspLeuProTyrrArgProAsp 220

```

```

DB  832 TTGCTGACCTCCAGCGGAGGTGCGGGCGCGCTTACGAGACGACCTGCCCTACCGCCGAC 891
QY  221 AlaValIleuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValIysAla 240
DB  892 GCCCTCTCTCTCTCTGCGCGCGCATCGACCTGTGACCGCGCGCCGAGTCAAGGCC 951
QY  241 ProArgHisIleValAlaIleProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
DB  952 CCGCGGACCAATGTGCGGCCCCCGGAGGACCAATACCGGAGCAGCCGACCTCC 1011
QY  261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
DB  1012 AGTGTCTTCTTCTCCGCTCCCAACGGGACCTTACCTTCTCCGTCGGCGCGCGGAG 1071
QY  281 CysGlyPheAspValSerLeuAspGlyGlnThrAlaThrIleGlnAspTrpIleGlyGly 300
DB  1072 TGCCTCTTCATGTACACCTCGGAGGAGAACCGCCACGTTCCAGGATTCGGGGGC 1131
QY  301 AsnTyrrValAsnIleAArgThrSerLyseAla 311
DB  1132 AACTACGTGAACATCGCGCCGCAATCCAGGCA 1164

RESULT 2
AAS17241
ID AAS17241 standard; DNA; 936 BP.
XX
XX AAS17241;
AC
XX
XX 12-MAR-2002 (first entry)
DT
XX
XX Streptomyces clavuligerus penicillin N expandase DNA sequence.
DE
XX
XX Penicillin N expandase; ring-expanding; penicillin G;
KM phenylacetyl-7-ADCA; amino desacetoxycephalosporanic acid; cephalosporin;
KM cephalixin; penicillin V; ds.
OS
XX
XX Streptomyces clavuligerus.
XX
XX Key Location/Qualifiers
FT CDS 1..936
FT /tag= a
FT /product= "Penicillin N expandase"
XX
XX MO2001.85951-A1.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-GB002047.
XX
XX 09-MAY-2000; 2000GB-00011185.
XX
XX (ACS-) ACS DOBPAR UK LTD.
XX
XX Johnson RI, Newbert RW;
XX
XX WPI; 2002-075247/10.
XX
XX P-PSDB; AAU11044.
XX
XX New enzyme useful for ring-expanding penicillin G to produce phenylacetyl
XX -7-amino desacetoxycephalosporanic acid, comprises penicillin expandase
XX having increased specificity for substrates such as penicillin G.
XX
XX Disclosure; Page 36-37; 42bp; English.
XX
XX The present invention relates to new penicillin expandases modified to
XX improve the ring-expanding activity for a substrate which is not the
XX natural substrate of the unmodified expandase. The invention is useful
XX for ring-expansion of penicillin G to produce phenylacetyl-7-ADCA (amino
XX desacetoxycephalosporanic acid) and the phenylacetyl side chain is
XX removed from phenylacetyl-7-ADCA to produce 7-ADCA, which is useful as a
XX starting point for the production of a range of semi-synthetic
XX cephalosporins, most notably cephalixin. The polynucleotide of the
XX invention is useful for producing a primer e.g. a PCR primer and as a

```


CC probe. The modification of the penicillin expandase enhances the activity
CC of the enzyme such as penicillin N expandase for penicillin G or V as a
CC substrate. The modified expandase has enhanced catalytic activity or
CC increased specificity for another substrate such as penicillin G. The
CC present nucleic acid sequence encodes the penicillin N expandase enzyme
CC of the invention

XX Sequence 936 BP; 154 A; 354 C; 284 G; 144 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.24e-167	Length:	936
Score:	1634.00	Matches:	310
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.68%	Mismatches:	0
Query Match:	99.88%	Indels:	0
DB:	6	Gaps:	0

10719236-1_232-1164 (1-311) x AAN92262 (1-936)

```
QY 1 MetAepThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHISGlnAsp 20
    |||||
DB 1 ATGGACACACGCGTGCCTCCACCTTCAGCTGCGCCGAATCCAGACGGCTTGACACCGAGAC 60
    |||||
QY 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
    |||||
DB 61 GAGTCCGCGAGGTGTCTGAGGACCAAGGCGCTTCTTATCTGACGACTGCGGTCTGACC 120
    |||||
QY 41 AapThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGlnHisGlySerGlu 60
    |||||
DB 121 GACACCGACCTGAAGTCGCGCAAGGACCTGTCATGACTTCTTGACGACGCGACGCGAG 180
    |||||
QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
    |||||
DB 181 GCGGAGAACGCGCGCTCCTGACCTGCGCGGCTCCACCATCGCGCGCTTCAACGCGGCTG 240
    |||||
QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
    |||||
DB 241 GAGTCCGAGACACCGCCCAAGATCAACATCCGAGACTACTCCGACATCTGATGTC 300
    |||||
QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGlnArgIleTyrThr 120
    |||||
DB 301 TACTCGATGGGACCGCGGACCAACCTCTTCCGCTCGGTGACTTCGAGCGGATCTGGAGC 360
    |||||
QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
    |||||
DB 361 CAGTACTTGACCGCGCAGTACACCGCTCCGCGCGGTGCGCGGAGGCTCTGCGGGG 420
    |||||
QY 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
    |||||
DB 421 ACCGGGACCGGACCGGCGCGGGGTGAGGCTTCTGACTGCGAGCCGCTGCTGCGG 480
    |||||
QY 161 PheArgTyrPheProGlnValProGlnHisArgSerAlaGluGlnProLeuArgMet 180
    |||||
DB 481 TTCGGTACTTCCCGGAGGTCCCGGAGCACCGGCGCGGAGGAGAGAGCCCTGCGGATG 540
    |||||
QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
    |||||
DB 541 GCGCGGACTACGACCTGTGATGATCACCCTCATTCACAGACACCTCGCGCAACGCG 600
    |||||
QY 201 PheValSerLeuGlnAlaGluValGlyAlaPheThrAspLeuProTyrArgProAsp 220
    |||||
DB 601 TTCGTACGCTCCAGGCGCGAGGTCCGCGCGGCTTTCACGAGACTGCGCTCCGAGAC 660
    |||||
QY 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValIleValysAla 240
    |||||
DB 661 GCGGTCTGCTCTTCTGCGGCGCATCGGACCTGTGACCGGCGGCGGCGGCAAGGCC 720
    |||||
QY 241 ProArgHisIleValAlaIleAlaProArgAspArgGlnIleAlaGlySerSerArgThrSer 260
    |||||
DB 721 CCGCGGACCATGTCCGCGCGCGCGCGGAGCACAGATGCGGCGGCGCGCACCTTC 780
    |||||
QY 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
    |||||
```

```
DB 781 AGTGTCTTCTCCCTCCGTCACAGCGGACCTTACACTTCTCCGTCGCGGCGCGGAG 840
    |||||
QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGly 300
    |||||
DB 841 TCGGCTTCTGATGTCTGACGCTGACGCGGACCGGACGATTCAGATTGATCGGGGCG 900
    |||||
QY 301 AsnTyrValAsnIleArgArgThrSerIysAla 311
    |||||
DB 901 AACTACGTGACATCCGCGGACATCCAGGCA 933
    |||||
```

RESULT 3

AAN92262

ID AAN92262 standard; DNA; 1230 BP.

XX AAN92262;

XX 25-MAR-2003 (revised)

XX 30-MAR-1990 (first entry)

XX Deacetoxycephalosporin C synthetase.

XX Deacetoxycephalosporin C synthetase; DAOCs; cephalosporin; penicillin N;

XX antibiotics.

XX Streptomyces clavuligerus.

XX Key Location/Qualifiers

XX mat_peptide 232..1167

XX FT /tag= a

XX FT /product= "DAOCs"

XX PN EP341892-A.

XX PD 15-NOV-1989.

XX PF 09-MAY-1988; 88US-00192273.

XX PR 09-MAY-1988; 88US-00192273.

XX PA (ELIT) LILLY & CO ELI.

XX PI Ingolia TD, Kovacevic S, Miller JR, Skatrud PL;

XX DR WPI: 1989-334231/46.

XX DR P-PSDB; AAP93215.

XX PT New recombinant DNA encoding de-acetoxy-cephalosporin C synthetase - for

XX PT increasing or inducing cephalosporin synthesis in microorganisms or in

XX PT vitro.

XX PS Claim 3; Page 27-28; 46pp; English.

XX CC The sequence encodes deacetoxycephalosporin C synthetase (DAOCs) which

XX CC catalyses expansion of penicillin N. Cephalosporium strains transformed

XX CC with vectors carrying the gene will produce antibiotics more efficiently.

XX CC (Updated on 25-MAR-2003 to correct PW field.) (Updated on 25-MAR-2003 to

XX CC correct DR field.)

XX SQ Sequence 1230 BP; 207 A; 455 C; 384 G; 184 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.87e-166	Length:	1230
Score:	1630.00	Matches:	310
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.63%	Indels:	0
DB:	1	Gaps:	0

10719236-1_232-1164 (1-311) x AAN92262 (1-1230)

```
QY 1 MetAepThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHISGlnAsp 20
    |||||
DB 232 ATGGACACACGCGTGCCTCCACCTTCAAGCTGCGGACCTCCAGACGCGCTGACACGAGAC 291
    |||||
```

```
OY 21 GIUPEHATGATGCGTLEUAATGAspIySGIyleuPheTYrleuThraSPcGlyleuThr 40
Db 292 GAGTTCGCCAGGTGCTTGAAGGACAAAGGCGCTCTTCTATCGACGCGACTGCGGTCTGACC 351
OY 41 AspThrgIuLeuLYSeSerAlalySAspIleValIleAspPhePheGlyNHISGlySerGlu 60
Db 352 GACACCGAGCTGACTCGGCGCAAGACATCGTCATCGACTTCTTGGACACGCGACGAG 411
OY 61 AlaGluYSarGAlaValIThrSerProValProThrMetArGArgGlyPheThrgIyleu 80
Db 412 GCGGAGAAAGCGCGCGTCACTCGCCCGTCCCAACATGCGCGCGGCTTCACCGGAGTGTG 471
OY 81 GluSerGusSerThraGlnIleThraSerThrgIySerTYrSerAspTYrSerMetCys 100
Db 472 GAGTGGAGAGACCGCCCAATATCACCAATACCGGAGCTACTCCGACTACTCGATGTGCG 531
OY 101 TYrSerMetGlyThraAlaAspAsnleuPheProSerGlyAspPheGluArgIleTyrThr 120
Db 532 TACTGATGGGACCGCGGACCAACTCTTCCCGTCCGGTGACTTGAACGGATCTGAGAC 591
OY 121 GlnTYrPheAspArgGlnTYrThraAlaSerArGAlaValAlaArgGluValleuArgAla 140
Db 592 CAGTACTTGCAGCGCCAGTACACCGCCCTCCCGCGGTGCGCCGAGAGTCTCGTCCGGGCG 651
OY 141 ThrgIyThrgIuProAspGlyGlyValGluAlaPheleuAspCysGluProleuAspArg 160
Db 652 ACCGGGACCGAGCCCGAGCGCGGGTTCAGAGCTTCTCGACTGCGAGCGCTGTGTGG 711
OY 161 PheArgTYrPheProGlnValProGluHISarGSerAlaGluGluGlnProleuArgMet 180
Db 712 TTCGGCTACTTCCCGCAGGTCTCCGAGACCGCGACGCGGAGACACCGCTCGCGGANG 771
OY 181 AlaProHISTYrAspLeuSerMetValThrleuIleGlnGlnThrProCysAlaAsnGly 200
Db 772 GCGCGGCACTACGACCTGTGATGTGTACCCCTCATCCAGCAGACACCTCGCGCCAAACGCG 831
OY 201 PheValSerleuGlnAlaGluValGlyAlaPheThraSPcGlyleuProTYrArgProAsp 220
Db 832 TTTCGACGCTTCCAGGCGCGAGGTCCCGCGGCTTTCAGACCTTCCCTTACCGTCCGAGC 891
OY 221 AlaValleuValPheCysGlyValIleAlaThrleuValThrgIyGlyValIleYsAla 240
Db 892 GCCGCTCTGTCTTTCGGGGCGCCANCGGACCTGTGTGACCGCGGCGGAGTCAAGGCC 951
OY 241 ProArGHSHisValAlaAlaProArGArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 952 CCCCCGCAACATGTGCGCGCGCCCCCGCAGGGACAGATAGCGGCGACGCGCACCTCC 1011
OY 261 SerValPhePheleuArProAsnAlaAspPheThraSPcGlyleuProleuAlaArgGlu 280
Db 1012 AGTGTGTCTTCTCTCGTCCCAACGCGGACTTCACTTCTCCGTCCGCTGGCGCGCAG 1071
OY 281 CysGlyPheAspValSerleuAspGlyGluThraIarThrPheGlnAspTYrIleGlyGly 300
Db 1072 TCGCGCTTCGATGTGAGCTGTGACGCGGACGCGCACGTTCCAGATTGATCGGGGCG 1131
OY 301 AsnTYrValAsnIleArGArgThrSerIysAla 311
Db 1132 AACTACGTGAACATCCGCGCACATCCAAAGGCA 1164
RESULT 4
AAQ88789
ID AAQ88789 standard, DNA; 939 BP.
XX
AC AAQ88789;
XX
DT 16-OCT--2003 (revised)
DT 25-MAR--2003 (revised)
DT 04-OCT--1995 (first entry)
XX
DE Nocardia lactamdurans expandase gene (cefe).
XX
```

```
KW Primer: amplified; PCR; expandase gene; cefe; Nocordia lactamdurans;
KW Streptomyces clavuligerus; expression cassette; acyltransferase; fungus;
KW Penicillium chrysogenum; hybrid promoter; Aspergillus nidulans; 7-ADCA;
KW 7-amino-desacetoxycapthalosporanic acid; cephalosporin; antibiotic; ds.
XX
OS Amycolatopsis lactamdurans.
XX
PN MO9504148-A1.
XX
PD 09-FEB-1995.
XX
PF 29-JUL-1994; 94WO-EP002543.
XX
PR 30-JUL-1993; 93EP-00202259.
PR 24-DEC-1993; 93BP-00203696.
XX
PA (KONN) GIST-BROCADES NV.
XX
PI Bovenberg RAL, Koekman BP, Hoekema A, Van Der Laan JM, Verweij J;
XX
DR WPI: 1995-082231/11.
XX
PT 7-amino-desacetoxy-cephalosporanic acid produ. in Penicillium chrysogenum
XX - by simultaneous expression of expandase and acyl-transferase.
XX
PS Example 1, Fig 8; 37pp; English.
XX
CC The nucleotide sequence of the Nocardia lactamdurans expandase gene cefe
CC as amplified by primer AAQ84983-4. The sequence differs from the
CC published sequence (Coque et al., Mol. Gen. Genet. 226 (1993), 453-458).
CC The amplified sequence lacks 3 bases: a G from pos. 120, a C from pos.
CC 122 and a G from pos. 124 of the above published sequence. This causes a
CC loss of the proline residue at amino acid pos. 41 of the corresponding
CC amino acid sequence. The expandase gene (cefe) from either Nocordia
CC lactamdurans or Streptomyces clavuligerus were amplified by PCR and
CC inserted into an expression cassette for simultaneous expression of the
CC cefe gene and the gene encoding an acyltransferase. The expression
CC of the cassette is placed in the fungus Penicillium chrysogenum. Expression
CC of the genes in the cassette is driven either by a trp-lac hybrid promoter
CC or the promoter from the Aspergillus nidulans gpdA gene. The terminator
CC is the 3'-end of the P.chrysogenum pombE gene (see AAQ84983-95). The
CC cassette is used in the production of 7-amino-desacetoxycapthalosporanic
CC acid (7-ADCA), an intermediate in the production of cephalosporin
CC antibiotics. Note: the sequences shown in this patent are identical to
CC those in patent WO 95/04149. (Updated on 25-MAR-2003 to correct PN
CC field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 939 BP; 153 A; 325 C; 326 G; 135 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5,57e-121 Length: 939
Score: 1210.00 Matches: 220
Percent Similarity: 83.87% Conservative: 40
Best Local Similarity: 70.97% Mismatches: 50
Query Match: 73.96% Indels: 0
DB: 2 Gaps: 0
10719236-1_232-1164 (1-311) x AAQ88789 (1-939)
OY 2 AspThrThraValProThraPheSerleuAlaGluLeuGlnGlnGlyleuNHISGlnAspGlu 21
Db 7 GACGCGACCGTGCAGACCTTGCATCTGCGCGAGCTGTGAGGAGCTTGCACGAGAGAG 66
OY 22 PheArGArgCysleuAArgAspIySGIyleuPheTYrleuThraSPcGlyleuThraSP 41
Db 67 TTCCGCACTGCTTCCCGAGAAAGCGGTGTCTTCACTCAAGGCGACCGGAGTGTCCGAG 126
OY 42 ThrgIuLeuLYSeSerAlalySAspIleValIleAspPhePheGlyNHISGlySerGluAla 61
Db 127 GCGGACCAACGCGCTGCGCGGAGATCGCGGTGACTTCTTGCACACGCGACCGAGGCC 186
OY 62 GluYSarGAlaValIThrSerProValProThrMetArGArgGlyPheThrgIyleuGlu 81
Db
```

```

Db      187 GAGAGAAAGCGGTGATGACGCCGATCCGACCATCCGGCGGGTACGCCGGGCTGGAG 246
Qy      82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyr 101
Db      247 TCCGAGAGACCGCGGAGATCAACGACCGGACGATCAACGACATCACTGAGTGTGAC 306
Qy      102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluIleTyrThrGln 121
Db      307 TCGATGGGACCGCGGACCAACCTGTTCCCGACGCCGAGTTGAGAAAGCGGTGGAGAGAC 366
Qy      122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
Db      367 TACTTCGCGCGGATGACCGCGCTTCGACGAGACGTGCGGGGAGGTGTGACCTCGATC 426
Qy      142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
Db      427 GGGCGGAAACCCGAGGTGCGGACGAGCGCTTCCTGACCTGCAACCCCTGCGGCCG 486
Qy      162 ArgTyrPheProGlnValProGluIleSarGserAlaGluGlnGlnProLeuArgMetAla 181
Db      487 CGCTACTTCCCGGAGGTGCGCGGAGATCGCGTGGCCGAGAGACGCCCTCGCGATGGCC 546
Qy      182 ProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPhe 201
Db      547 CCGCACTACGACCTCTCGATCGTCACCTCGATCCACCAAGCCCTTGCGGAAACGGGTTTC 606
Qy      202 ValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAla 221
Db      607 GTCAGCGCTCAGGTCGAGGTGAGCGGGTCTATGTGACATCCCGCGGACCGCGGCGG 666
Qy      222 ValLeuValPheCysGlyAlaIleAlaThrIleValThrGlyGlyGlnValIysAlaPro 241
Db      667 GTGCTGTGTCTTCGCGCGCGGAGTGGCGGACCTGTGTGGCCGACCGCGCATCAAGGGGCC 726
Qy      242 ArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
Db      727 AAGACACACGCTGGCGCGCGCGCGGACGAGCAAGCGGTGGGACGACCGCACCTTCACG 786
Qy      262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCys 281
Db      787 GGTGTTCTTCCTGCGCGCGGACGCGGACTTCGCTTCGCTGCGCGCGGCGACGGGAGTGC 846
Qy      282 GlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyIysAsn 301
Db      847 GGGTTGCACTGACGATCCCGCGCGGACCGCACCTTCGACGACTCGATCGCGGCAAC 906
Qy      302 TyrValAsnIleArgArgThrSerIysAla 311
Db      907 TACATCAACATCCGGAAGACCGCGCGCGC 936

```

RESULT 5
AA082721
ID AA082721 standard; DNA; 939 BP.

AC AA082721;
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-OCT-1995 (first entry)
XX
DE N. lactamdurans cefE gene PCR 1 product.
XX
XX N. lactamdurans; P. chrysogenum; cefE gene; PCR 1 product;
KW expression cassette; 7-aminodeacetoxy-cephalosporanic acid;
KW expandase gene; cephalosporin antibiotics; ss.
XX
XX Amycolatopsis lactamdurans.
XX
XX WO9504149-A1.
XX
XX 09-FEB-1995.
XX
XX 29-JUL-1994; 94WO-EP002544.
PF

```

XX      30-JUL-1993; 93BP-00202260.
PR      24-DEC-1993; 93BP-00203695.
XX      (KONN ) GIST-BROCADES NV.
XX      Bovenberg RAL, Koekman BP, Hoekema A, Van Der Laan JM, Verweij J;
XX      WPI, 1995-082232/11.
XX      7-amino-desacetoxy-cephalosporanic acid prodn. in Penicillium chrysogenum
XX      PT - transformed with expandase gene, using 3,3'-thiodi:propionic acid as
XX      side chain precursor and deacylation of intermediate.
XX      PS Example 1, Fig 8; 37bp; English.
XX      CC AA082720 is the N. lactamdurans cefE gene, it is compared with AA082721
XX      CC the PCR 1 generated N. lactamdurans cefE (expandase) gene were constructed
XX      CC cassettes for the N. lactamdurans cefE (expandase) gene can now be used
XX      CC using the PCR 1 product. The transformed P. chrysogenum can now be used
XX      CC for 7-amino-desacetoxy-cephalosporanic acid prodn. an intermediate for
XX      CC cephalosporin antibiotics. (Updated on 25-MAR-2003 to correct PN field.)
XX      CC (Updated on 16-OCT-2003 to standardise OS field)
SQ      Sequence 939 BP; 153 A; 325 C; 326 G; 135 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,57e-121 Length: 939
Score: 1210.00 Matches: 220
Percent Similarity: 83.87% Conservative: 40
Best Local Similarity: 70.97% Mismatches: 50
Query Match: 73.96% Indels: 0
DB: Gaps: 0

10719236-1_232-1164 (1-311) x AA082721 (1-939)
Qy      2 AspThrTrpValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGlu 21
Db      7 GACGGACCGTGCCGACCTTCGATTTGGCCGAGCTGCTGAGGGCTTCACACGAGAGAG 66
Qy      22 PheArgArgCysLeuArgAspGlyGlyLeuPheTyrLeuThrAspCysGlyLeuThrAsp 41
Db      67 TTCGCGCACCTGCTGCGGAGAGAGGCGGTCTTCACTCAAGGGGACCGGGCTGCCGAG 126
Qy      42 ThrGluLeuIysSerAlaIysAspIleValIleAspPhePheGlnHisGlySerGluAla 61
Db      127 GCGGACCAAGCGCTCGCGCGGAGATCCGCGGACTTCTTCGACCAAGGACCGAGGCC 186
Qy      62 GluIysAspAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlu 81
Db      187 GAGAAAGAGCGGTGATGACGCGCCGATCCGACCATCCGCGCGGGGTACCGCGGCTGAG 246
Qy      82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyr 101
Db      247 TCCGAGAGACCGCGGACATCAACGACCGGACGATCAACGACTACTCGATGTCGATC 306
Qy      102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluIleTyrThrGln 121
Db      307 TCGATGGGACCGCGGACCAACCTGTTCCCGACGCCGAGTTGAGAAAGCGGTGGAGAGAC 366
Qy      122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
Db      367 TACTTCGCGCGGATGACCGCGCTTCGACGAGACGTGCGGGGAGGTGTGACCTCGATC 426
Qy      142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
Db      427 GGGCGGAAACCCGAGGTGCGGACGAGCGCTTCCTGACCTGCAACCCCTGCGGCCG 486
Qy      162 ArgTyrPheProGlnValProGluIleSarGserAlaGluGlnGlnProLeuArgMetAla 181
Db      487 CGCTACTTCCCGGAGGTGCGCGGAGATCGCGTGGCCGAGAGACGCCCTCGCGATGGCC 546
Qy      182 ProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPhe 201

```

```
Db      CCGACCTACGACCTTCATGCTACCTGATCCACGAGACCCCTTGCGGCAACGGCTTC 606
Qy      202 ValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeuProTyrArgProAspAla 221
Db      607 GTACAGCTCGTACAGTGCAGGTGAGCTGATGTGACATCCCGGCGCACCGCGCGCG 666
Qy      222 ValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValValAlaPro 241
Db      667 GTGCTGTGTTCTGCGGCGGCTGCGAGCTGGCGGAGCGGCGGCGATCAAGCGGCC 726
Qy      242 ArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
Db      727 AAGACCCAGCTGCGCGCGCGCGCGCGGACCAAGCGGAGGAGAGCGACCTCCAGC 786
Qy      262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCys 281
Db      787 GTGTTCTTCTGCGCGCCCAACGGGGAATTCCGCTTCTCGTGCCGCGGCGCAGGAGTGC 846
Qy      282 GlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGlyAsn 301
Db      847 GGGTTCGACGTACGATCCCGCGCGGACCGGACCTTCGACGACTGATCGCGCGGCAAC 906
Qy      302 TyrValAsnIleArgArgThrSerIysAla 311
Db      907 TACATCAACATCCGGAAGACCGCGCGCGC 936
RESULT 6
AA082720
ID      AA082720 standard; DNA; 942 BP.
AC      AA082720;
XX      16-OCT-2003 (revised)
DT      25-MAR-2003 (revised)
DT      04-OCT-1995 (first entry)
XX      N. lactamdurans cefb gene.
DE      N. lactamdurans cefb gene.
XX      N. lactamdurans; P. chrysogenum; cefb gene; expression cassette;
KM      7-amino-desacetoxy-cephalosporanic acid; expandase gene;
KW      cephalosporin antibiotics; ss.
XX      Amycolatopsis lactamdurans.
XX      MO9504149-A1.
XX      09-FEB-1995.
XX      29-JUL-1994; 94WO-BP002544.
XX      30-JUL-1993; 93BP-00202260.
XX      24-DEC-1993; 93BP-00203695.
XX      (KONN ) GIST-BROCADES NV.
XX      PA
XX      Bovenberg RAL, Koekman BP, Hoekema A, Van Der Laan JM, Verweij J;
PI      WPI; 1995-082232/11.
XX      7-amino-desacetoxy-cephalosporanic acid prodn. in Penicillium chrysogenum
PT      -transformed with expandase gene, using 3'-thiodi:propionic acid as
PR      side chain precursor and deacylation of intermediate.
XX      Example 1; Fig 8; 37pp; English.
XX      AA082720 is the N. lactamdurans cefb gene, it is compared with AA082721,
CC      the PCR 1 generated N. lactamdurans cefb gene. P. chrysogenum expression
CC      cassettes for the N. lactamdurans cefb (expandase) gene were constructed
CC      using the PCR 1 product. The transformed P. chrysogenum can now be used
CC      for 7-amino-desacetoxy-cephalosporanic acid prodn. an intermediate for
CC      cephalosporin antibiotics. (Updated on 25-MAR-2003 to correct PN field.)
CC      (Updated on 16-OCT-2003 to standardise OS field)
```

```
XX      SQ      Sequence 942 BP: 153 A; 326 C; 328 G; 135 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 7.66e-120 Length: 942
Score: 1199.50 Matches: 220
Percent Similarity: 83.60% Conservative: 40
Best Local Similarity: 70.74% Mismatches: 50
Query Match: 73.32% Indels: 1
DB: 2 Gaps: 1
10719236-1_232-1164 (1-311) x AA082720 (1-942)
Qy      2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGlu 21
Db      7 GACCGCACCTGCGGACCTTCGATCTGGCCGAGCTGCTAGAGGCTTGACACCGAGAG 66
Qy      22 PheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeu---Thr 40
Db      67 TTCGGCACCTGCGGCGCGGAGGAGGCGTGTCTACTCAAGGCGCACCGGCGTGGCGCC 126
Qy      41 AspThrGluLeuIysSerAlaLysAspIleValIleAspPheGlnHisGlySerGlu 60
Db      127 GAGCGGACCAACGCGCTCGCGCGGAGATCGCGTGGACTTCTTCACACCGCACCGCAG 186
Qy      61 AlAGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db      187 GCCGAGAAAGAGGGGTGATGACGCGCATCCCAACATCCGCGCGGATACGCGGAGCTG 246
Qy      81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db      247 GAGTCCGAGAGCACCGCGGACGATCAGCAACACCGGCAAGTACACCGACTCTCGATGTCG 306
Qy      101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGlnValGlyIleTrpThr 120
Db      307 TACTCGATGGGACACCGGGAACACTGTTCCAGCGCCAGTTCGAGAAAGCGGTGGAG 366
Qy      121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db      367 GACTACTTCCGCGGAGATGACCGCGCTTCCGAGAGCTCGCGGCGAGTGTGACTCTCG 426
Qy      141 ThrGlyThrGluProAspArgGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db      427 GTCGGCGCGGAACCGGAGTGGGATGAGCGCTTCTTCGACGTGCGAACCCCTGCTCGC 486
Qy      161 PheArgTyrPheProGlnValProGlnHisArgSerAlaGlnGlnPheLeuAspMet 180
Db      487 CTGGCTACTTCCCGAGGTGCGCGAGATCGGTGGCCGAGAGCGCGCTGGATG 546
Qy      181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db      547 GCCCGGACTACGACCTTCGATGCTACCTGATCCACGACACCTTGCGGGAACGGG 606
Qy      201 PheValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeuProTyrArgProAsp 220
Db      607 TTGCTCAGCTCGAGGTGCGAGGTGAGCGGTCTATGTGACATCCCGGCGGACCGGCG 666
Qy      221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValValAla 240
Db      667 GCGGTGTGTTGTTCTGCGCGCGGTGCGACGCTGTGGCCGACGCGGATCAAGCGG 726
Qy      241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db      727 CCCAAGCACACGTGGCGCGCGCGCGGAGACAGCGGAGTGGGACGACCCGACCTCC 786
Qy      261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db      787 AGCGTGTCTTCTGCGCGCCCAACGGGGACTTCTCGGTGCGCGGCGCAGGAG 846
Qy      281 CysGlyThrAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300
Db      847 TGCGGTTCGACGTACGATCCCGCGCGGACCGGCACTTTCGACGACTGATCGCGGCG 906
```

QY 301 AsnTYrValAsnIleArgArgThrSerLYsAla 311
|||:::|||||:::|||||
Db 907 AACTACATCAACATCCGGAAGACCCGCCGCC 939

RESULT 7
AAT08693
ID AAT08693 standard; cDNA; 2672 BP.
XX
XX AAT08693;
XX
DT 16-OCT-2003 (revised)
DT 25-AUG-1996 (first entry)
DE Cephamycin biosynthetic enzyme gene cluster.
KW Cephamycin; cephalosporin; antibiotic;
KW C-7 hydroxycephem methyltransferase; 3'-methylcephem hydroxylase;
KW 3'-hydroxymethylcephem O-carbamoyltransferase; ss.
XX
XX Amycolatopsis lactamdurans; strain LC411.
OS
FH Key Location/Qualifiers
FT CDS 1..70
FT FT /*tag= a
FT FT /label= pcbc
FT FT /note= "3' region of pcbc gene"
FT RBS 73..80
FT FT /*tag= b
FT CDS 84..794
FT FT /*tag= c
FT FT /label= ORF7
FT FT /product= "C-7 hydroxycephem methyltransferase"
FT FT /note= "(Claim 8, page 40)"
FT misc_difference 504..602
FT FT /*tag= d
FT FT /note= "translated sequence of bases 503-602 given in the
FT FT specification do not correspond to the Genetic Code"
FT RBS 785..789
FT FT /*tag= e
FT FT /note= "putative RBS for ORF7/ORF8 cotranslation"
FT CDS 802..1668
FT FT /*tag= f
FT FT /label= ORF8
FT FT /product= "coupling protein"
FT FT /note= "(Claim 7, page 39-40)"
FT RBS 1678..1685
FT FT /*tag= g
FT FT /label= 9
FT CDS 1692..2624
FT FT /*tag= h
FT FT /label= ORF9
FT FT /product= "3'-methylcephem-hydroxylase"
FT FT /note= "cef gene (Claim 5, page 39)"
FT FT 2634..2667
FT FT /*tag= i
FT FT /note= "stem-loop may correspond to a transcription
FT terminator"
XX
XX WO9529253-A1.
XX
XX 02-NOV-1995.
XX PD
XX 17-APR-1995; 95WO-US004801.
XX PF
XX 22-APR-1994; 94US-00233605.
XX PR
XX (MERI) MERCK & CO INC.
XX PA
XX Coque UJR, Enguita FU, Fuente JL, Llaarena FU, Liras P, Martin JF;
XX PI
XX WPI, 1995-382998/49.
XX DR P-PSDB; AAR92151, AAR92152, AAR92153, AAR92154.
XX
XX DNA encoding late enzymes involved in cephamycin biosynthesis - and

PT related proteins, used to transform cells for cephamycin prodn.
XX
XX Disclosure; Fig 3A-C; 61pp; English.
XX
CC An internal sequence (AAT08693) of a 5.4 kb BamHI cDNA fragment of the
CC Nocardia lactamdurans cephamycin C gene cluster was isolated from a phage
CC clone using a cefE gene probe. It contains the 3' region of the pcbc
CC gene, open reading frames (ORF7 and ORF8) that code for a 2-component
CC system comprising C-7 hydroxylase methyltransferase (AAR92152) and a
CC coupling protein (AAR92153), and the cefE gene (ORF9) encoding 3'-
CC methylcephem hydroxylase (AAR92154). Another isolated cDNA (AAT08694)
CC included the downstream cmcH gene (ORF10) coding for 3'-
CC hydroxymethylcephem O-carbamoyltransferase (AAR92156). These sequences
CC can be used for prodn. of recombinant enzymes useful for industrial-scale
CC cephamycin synthesis. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 2672 BP; 441 A; 952 C; 901 G; 378 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,2e-94 Length: 2672
Score: 972.50 Matches: 187
Percent Similarity: 73.68% Conservative: 37
Best Local Similarity: 61.51% Mismatches: 79
Query Match: 59.44% Indels: 1
DB: 2 Gaps: 1
10719236-1_232-1164 (1-311) x AAT08693 (1-2672)
QY 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAArgIu 21
|||:::|||||:::|||||
Db 1698 GACAGACGGTACCGGCTTCCGATGAGTCCGACGCGGACGCGTCCGCGACGAG 1757
QY 22 PheArgArgCysLeuArgArgPheGlyLeuPheTYrLeuThrAArgCysGlyLeuThrAsp 41
|||:::|||||:::|||||
Db 1758 TTCGCGAGTGGGCGCGCG---CGCGGGCTTCTTACCTCCACCGGGGTACGGCGCCCA 1814
QY 42 ThrGluLeuYserAlaLeuAspIleValIleAspPheGlnHisGlySerGluA 61
|||:::|||||:::|||||
Db 1815 CGAGACCGCGGGTGGCCGACCGACCGCGGATGAGCTTCTTGGCCCAAGGCGCGCGAG 1874
QY 62 GluLYsArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlu 81
|||:::|||||:::|||||
Db 1875 GAGAGACGAGCGCGTACACAGAGTCCGACGACGCGCGGGGTATCGCGCGGTGAG 1934
QY 82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTYrSerAspTYrSerMetCysTYr 101
|||:::|||||:::|||||
Db 1935 GCGGAAAGCACCGCGCCAGGTACCAACACCGCGACCTACCGACTCATGTCTGTAC 1994
QY 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleThrProGln 121
|||:::|||||:::|||||
Db 1995 TCGATGGGCATCGCGCGGACCTGTCCCGTGAAGAGTTGAGTCCGTGACGAGAC 2054
QY 122 TyrPheAspArgGlnTYrThrAlaSerArgAlaValAlaArgGluValLeuAArgAlaThr 141
|||:::|||||:::|||||
Db 2055 TACTTCGACAGCGCTGTACCGCGCGCGGAGACCGCGCGCGTGTCTGTACCGCGCG 2114
QY 142 GlyThrGluProAspArgGlyValGlnAlaPheLeuAspCysGluProLeuLeuAArgPhe 161
|||:::|||||:::|||||
Db 2115 GGCACCTACGACCGCGGAGGACCTGTGACACCTGTCTGTACCTGTGACCGCGTGTGCGCTG 2174
QY 162 ArgTYrPheProGlnValProGlnHisArgSerAlaGluGlnInProLeuAArgMetAla 181
|||:::|||||:::|||||
Db 2175 CGGTACTTCCCGAGAGTCCCGAGAGCACCGCGCGCGGAGTACGACCGCGCGAGTGGCC 2234
QY 182 ProHisTYrAspLeuSerMetValThrLeuIleGlnInThrProCysAlaAsnGlyPhe 201
|||:::|||||:::|||||
Db 2235 CGGACCTACGACCGTGTCTGTATCATTCACGACGACCGCGTGTGCGGACGAGTTTC 2294
QY 202 ValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeuProTYrArgProAspAla 221
|||:::|||||:::|||||
Db 2295 GTCAAGCTGACGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2354
QY 222 ValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValAlaPro 241

```
DB 2355 GTGGTGGTCTGTGGCGCGCATCGCGCCCTGGTACCCAGGCGCGGCGCCGCCGCC 2414
OY 242 ATGHSIAISVALAIALAProArgArgAspGlnIlealagIySerSerArgThSerSer 261
DB 2415 AACCCACCACTGTCTCCCCCGGACCGAGCATGCTCAAGGGCAGCAGCCGACCTCGAGC 2474
OY 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCys 281
DB 2475 GTGTTCTTCTGCGCCGTCGACCGATTTCATCTTCTGCTGCCGACGCCAGGAAGTAC 2534
OY 282 G1PheAspValSerLeuAspG1yG1uThrAlaThrPheGlnAspTrpIleG1yG1yAsn 301
DB 2535 GCCCTCGACGTAGCCTGGACATGAGAGAGGACCTTCGGGACTGATCGGAGCAAC 2594
OY 302 TyrValAsnIle 305
DB 2595 TACGTCACGATG 2606

RESULT 8
AAT00591
ID AAT00591 standard; cDNA, 972 BP.
XX
XX AAT00591;
XX AC
XX 16-OCT-2003 (revised)
DT 13-APR-1996 (first entry)
XX
DE cefF (ORF9) encoding 3'-methylcephemhydroxylase.
KW cephamycin C biosynthesis; cephem-carbamoyltransferase; hydroxylase;
KW methyltransferase; cluster; cmch; cmcl; cmcl; cefF; sf.
OS Amycolatopsis lactamdurans.
XX
XX Key Location/Qualifiers
XX CDS 1..924
XX FT /*tag= a
XX FT /*note= "3'-methylcephemhydroxylase"
XX FT /*tag= b
XX FT /*note= "putative terminator"
XX
XX MO9529253-A1.
XX
XX 02-NOV-1995.
XX
XX 17-APR-1995; 95WO-US004801.
XX
XX 22-APR-1994; 94US-00233605.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Coque JJR, Enguita FJ, Fuente JI, Llaarena FJ, Liras P, Martin JF;
XX
XX MPI; 1995-382998/49.
XX
XX P-PSDB; AAR84175.
XX
XX DNA encoding late enzymes involved in cephamycin biosynthesis - and
XX related proteins, used to transform cells for cephamycin prodn.
XX
XX Claim 5; Page 39; 61pp; English.
XX
XX Three genes (AAT00591-93) located in the cluster of cephamycin C
XX biosynthesis in N. lactamdurans which encode deacetoxycephalosporin C
XX hydroxylase and two other proteins, which introduce the methoxyl group at
XX C-7 have been isolated and sequenced. The sequence of one of the latter
XX proteins resembles both cholesterol hydroxylases and methyltransferases
XX of different origins acting on hydroxyl groups present in aromatic or
XX quirono-type compounds; both proteins are required for the hydroxylation
XX at C-7 and the transfer of the methyl group from S-adenosylmethionine to
XX the 7-hydroxycephem intermediate. In addition, the isolation, nucleotide
XX sequence, and the characterization of a gene (AAT00590) encoding a 3'-
```

```
CC hydroxymethylcephem O-carbamoyl- transferase, designated cmch has been
CC shown. (updated on 16-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 972 BP, 165 A; 358 C; 314 G; 135 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.02e-94 Length: 972
XX Score: 967.50 Matches: 186
XX Percent Similarity: 73.84% Conservative: 37
XX Best Local Similarity: 61.59% Mismatches: 78
XX Query Match: 59.14% Indels: 1
XX DB: 2 Gaps: 1
XX
XX 10719236-1_232-1164 (1-311) x AAT00591 (1-972)
OY 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlnIleuHisGlnAspGluPheArg 23
DB 4 ACGGTACCGGTCTTCAGCATGAGCGCAACTCGCGACCGGCTCGCGCAGGACGATTCGCG 63
OY 24 ArgCysLeuArgAspGlyLeuPheTyrLeuThrAspCysGlyLeuThrAspThrGlu 43
DB 64 GAGTGGGCCCGC--CGCGGGTCTTCTTACTCTACCGGTTACCGGCCGACGACGAC 120
OY 44 LeuYSerAlaYsaAspIleValIleAspPheGlnIleGlySerGluAlaGluYs 63
DB 121 CACCGGGTGGCCACCGACACCGGAGTGACTTCTTGCCCAAGGACGCGCGAGAGAG 180
OY 64 ArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGluSerGlu 83
DB 181 CAGCCCGTGACCAAGAGTTCGACCATCGCGCGGGTACTCGCGGCTGGAGGCGGA 240
OY 84 SerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyrSerMet 103
DB 241 AGCACCGCCAGGTGACCAACCGGACCTTACCTACCGACTTCTCATGTGTCTGATG 300
OY 104 G1YThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTrpThrGlnTyrPhe 123
DB 301 GGATCGGCGCAACCTGTTCCGCTGAGAGAGTTCAGTTCGATGAGACGACTACTTC 360
OY 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThrGlyThr 143
DB 361 GACAGCCTGTACCGCGCGGCGAGAGACCGCGCGCTGTGTGACCGCGCGGCGAC 420
OY 144 GluProAspGlyGluValGluAlaPheLeuAspCysGluProLeuLeuArgPheArgTyr 163
DB 421 TACAGCGCGAGACCTCGACACCTCTCGACTCGACACCGGCTGCTGCGCTGAC 480
OY 164 PheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAlaProHis 183
DB 481 TTCCGGAGGTCGCGAGCACCGCGCGCGAGTACAGAGCACGCGGATGCGCCGAC 540
OY 184 TyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysValAlaGlnIlePheValSer 203
DB 541 TACGACCTGTCCACACCTTCATCCACCAACCCCGGCGCAAGGATTCGTGAGC 600
OY 204 LeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAlaValLeu 223
DB 601 CTGACGCGCAAGTGCAGCGGTGAGTGAAGCTGCGCGACGATCCAGACGCGGTGTC 660
OY 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValYsaAlaProArgHis 243
DB 661 GTGCTGTGCGCGGATCGCGCGCTGTGTACCAAGGCGCGGTGCGCGCGCAACAC 720
OY 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
DB 721 CACGTGTGTTCCTCCGAGCGGACGATGCTCAAGGCGACGACGACCTTGACCGTGTTC 780
OY 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283
DB 781 TTCTTGGCGCGCGTGCACCGATTTCCTTCTGCTGCCAGCGCGAGAGTACGGCTTC 840
OY 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGlyAsnTyrVal 303
```

Db 841 GACGTGAGCTGGACATGGAAGCGACCTTCGGCGACTGGATCGGACCAACTACGTC 900
QY 304 Asn116 305
Db 901 ACCGATG 906
RESULT 9
AAQ10190
ID AAQ10190 standard; DNA; 23666 BP.
XX
AC AAQ10190;
XX
DT 25-MAR-2003 (revised)
DT 27-MAR-1991 (first entry)
XX
DE Cephalosporin antibiotic biosynthetic genes.
XX
KW cephalosporin; antibiotic; S-(L-alpha-aminoadipyl)-L-cysteinyL-D-;
KW valine synthetase; isopenicillin N synthetase; isopenicillin N epimerase;
KW deacetoxycephalosporin C synthetase; beta-lactamase;
KW deacetoxycephalosporin C hydrolase; ss.
OS Lysobacter lactamgenus.
XX
FH Key Location/Qualifiers
FT CDS 6..2819
FT FT /tag= 1
FT FT /label= ORF 9
FT CDS 2887..4200
FT FT /*tag= h
FT FT /label= ORF 8
FT FT 4211..5443
FT FT /*tag= g
FT FT /label= ORF 7
FT FT 5524..16692
FT FT /*tag= a
FT FT /label= ORF 1
FT FT 16761..17741
FT FT /*tag= b
FT FT /label= ORF 2
FT CDS 17802..18761
FT FT /*tag= c
FT FT /label= ORF 3
FT CDS 18798..19739
FT FT /*tag= d
FT FT /label= ORF 4
FT CDS 19802..21061
FT FT /*tag= e
FT FT /label= ORF 5
FT CDS 21186..22343
FT FT /*tag= f
FT FT /label= ORF 6
XX
XX JP02291274-A.
XX
XX PD 03-DEC-1990.
XX
XX PF 10-JAN-1990; 90JP-00003762.
XX
XX PR 01-FEB-1989; 89JP-00024710.
XX
XX PA (TAKE) TAKEDA CHEM IND LTD.
XX
XX WPI; 1991-018854/03.
XX DR P-PSDB; AARI0145, AARI0688, AARI0689, AARI0690, AARI0691, AARI0692,
XX AARI0693, AARI0694, AARI0695.
XX
XX PT Prep. of cephalosporin series antibiotics - comprises culturing
XX transformant of microbe transformed by plasmid contg. new DNA fragment.
XX
XX PS Claim 4; Fig 1; 67p; Japanese.
XX
XX A fragment of the sequence comprising at least one of the cephalosporin

CC biosynthetic enzymes listed in the KEYWORDS can be cloned in a plasmid
CC and used to transform microbes, such as bacteria or yeast. Although DNA
CC is preferably isolated from L. lactamgenus, similar sequences could be
CC obtained from other bacteria containing genes coding for biosynthesis of
CC cephalosporin series antibiotics. See also AAQ10191-2. (Updated on 25-MAR
CC -2003 to correct PA field.)
XX
SQ Sequence 23666 BP; 3952 A; 8522 C; 7603 G; 3589 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2,996-90 Length: 23666
Score: 944.00 Matches: 178
Percent Similarity: 73.11% Conserved: 45
Best Local Similarity: 58.36% Mismatches: 82
Query Match: 57.70% Indels: 0
DB: 2 Gaps: 0
10719236-1_232-1164 (1-311) x AAQ10190 (1-23666)
QY 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlyAspGlu 21
Db 17808 GATTGAGCATCCAGATATTGACCTGGACCACTCGACACGGCGTCCGCTGATTCG 17867
QY 22 PheArgArgCysLeuArgAspGlyLeuPheTyrLeuThrAspCysGlyLeuThrAsp 41
Db 17868 TTCCGCAATCCCTCTTTCGAGCGCGCGCTTTTACGTCGCGGAGAGATTCGATCAA 17927
QY 42 ThrGluLeuLysSerAlaLysAspIleValIleAspPheGlnHisGlySerGluAla 61
Db 17928 ACCGAGCAGCGCAAGCGATGACCGCTCATGACCTGTTCGAGAACGGCAGCGGAA 17987
QY 62 GluLysArgAlaValThrSerProValProThrPheArgGlyPheThrGlyLeuGlu 81
Db 17988 CAGAGAACGGCTCAGGACCTGACCTGCAACGTCGCGCGGTTTTCGACCTGAG 18047
QY 82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerMetCysTyr 101
Db 18048 GCCGAGACAGCGCGCGATCACCAGGCGCGCAATACAGCACTACTCATGATGCTAT 18107
QY 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGlnHisGlyIleThrGln 121
Db 18108 TCGATCGCGCTGACGAGCAACCTGTTCCGCGCGGCTTTCGAAAGCCATCTGACCGGC 18167
QY 122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
Db 18168 TATTTCGACCGTTTCTACGCTCGCACCCGAGCAATCGCGCTTCGCTCGCGCTTC 18227
QY 142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
Db 18228 GACCGCGGCTCGACGAGAGCGTGAGCAAGTTCTTCGACTCGCATCGCTCGGTTTC 18287
QY 162 ArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAla 181
Db 18288 CGCTTTTCCCGAAGTCCCGAAGACCGCTGCGCGCGAGCGCGCGCGCATGCGG 18347
QY 182 ProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAspGlyPhe 201
Db 18348 CGCATTTACGACCTGTCATATCATACCAACATCCAGCAGAACCTTGGCCACCGCTTC 18407
QY 202 ValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeuProTyrArgProAspAla 221
Db 18408 GTCAAGCTGACAGTCGACAGTGCAGCGCCGTTACGTCATCTGCCGCGCGCGCGCTGC 18467
QY 222 ValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLysAlaPro 241
Db 18468 ATGTGTGTCTCTCGCGCGCGCTGCGCGCTGTTTCGCGCGCGAGGATCAAGCGCGG 18527
QY 242 ArgHisValAlaIleAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
Db 18528 CGCCACGAGTCGCTCGCTCCGCTCATGACGAGCGCGCTCGCAGAGAGCGCACTCCAGC 18587
QY 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCys 281

Dd		18588	GTGTTCTTTCCTCGGGCCGAAGCCGACTTCGGTTCGGTGCCGCCCGGCACAC	18647
Oy		282	GlyPheAspValSerLeuAspGlyIleThrAlaThr-PheGlnAspTrpIleGlyIleAsn	301
Dd		18648	GGATCTCATGTCGATTTCACCGGGCACAACCGCACCACTTCGGCGAAGTGATCGCGGCAAC	18707
Oy		302	TyrValasniLeArg	306
Dd		18708	TACGTCAACCTCGCG	18722
		RESULT 10		
		AAN81128		
ID		AAN81128	standard; DNA; 996 BP.	
XX				
AC		AAN81128;		
XX				
DT		24-OCT-2003	(revised)	
DT		25-MAR-2003	(revised)	
DT		29-OCT-1990	(first entry)	
XX				
DE		Deacetoxycephalosporin-C-synthetase and deacetylcephalosporin-C-		
DE		synthetase structural gene.		
XX				
KW		Deacetoxycephalosporin-C-synthetase; Deacetylcephalosporin-C-synthetase;		
XX		cephalosporin; DS DNA.		
OS		Acremonium chrysogenum.		
PN		EP281391-A.		
PD		07-SEP-1988.		
XX				
Pf		03-MAR-1988;	88EP-00301845.	
XX				
PR		04-MAR-1987;	87US-00021836.	
XX				
PA		(ELIL) LILLY & CO ELI.		
PI				
XX		Ingolia TD, Queener SW, Samson SM, Skatrud PL;		
DR		WPI; 1988-251947/36.		
PT		Recombinant DNA from Cephalosporium acremonium - encoding		
PT		deacetylcephalosporin C synthetase - di:acetox:cephalosporin C		
XX		synthetase activity.		
PS				
XX		Disclosure; Page ?; 82pp; English.		
CC				
CC		DNA sequences encoding deacetoxycephalosporin-C-synthetase and		
CC		deacetylcephalosporin-C-synthetase are used to construct vectors for a		
CC		Cephalosporium. Penicillium transformants may be used for cephalosporin		
CC		production. See also AAP80534, AAN81125, AAN81126, AAN81127 (Updated on		
CC		25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise		
OS		field)		
SQ		Sequence 996 BP; 176 A; 345 C; 310 G; 165 T; 0 U; 0 Other;		
		Alignment Scores:		
		Pred. No.: 5.05e-91 Length: 996		
		Score: 933.50 Matches: 178		
		Percent Similarity: 70.83% Conservative: 43		
		Best Local Similarity: 57.05% Mismatches: 90		
		Query Match: 57.06% Indels: 1		
		DB: 1 Gaps: 1		
		10719236-1_232-1164 (1-311) x AAN81128 (1-996)		
Oy		1 MetAspTrpThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAsp		20
Dd		1 ATGACTTCGAAAGGTCCCGCTCTTGTCGACACACTCAAGAAGCGCAAGTCTTCACC		60
Oy		21 GluPheArGArGCYsLeuArGAspLySgIyeuPhetyrIeuThraSpCYsgIyeuthr		40

Db	61	AGCTCCGCCGAGGCGCGTCACCAAGGATATCTTCTACTGACCGAAGGCGCTGTC	120
Qy	41	AspThrGluLeuLeuSerAlaLeuAspIleValIleAspPheGluHisGlySerGlu	60
Db	121	GACGACGACCAACACTCGCGCGCGTGAGACGTCGGTTCCTTTTCAGAAACGGAAGCGAG	180
Qy	61	AlaGluLeuSerAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu	80
Db	181	GAGGAGAAAGAGGCGCGTACCGCTGCGGACCGGTACCGCCCGCGGCGCTTCTGCGCTTC	240
Qy	81	GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys	100
Db	241	GAGTGGAGAGACACCGCGTGTACCGAGACGGGCAAGTACTCGACTACGACGTGC	300
Qy	101	TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleThrPhe	120
Db	301	TACTCCAGGGCGATCGGCGGCAACTGTTTCCGAAACCGGGCTTCGAGAGCGTCTGGAG	360
Qy	121	GlnTyrPheAspArgGlnTyrThrIleSerArgAlaValAlaArgGluValIleuArgAla	140
Db	361	GACTACTTCGACCGCATGTACGCGGACCGACCAAGATGTCTGGCGCGCGCTTCTCAACTCT	420
Qy	141	ThrGlyThrGluProAspGly--GlyValGluAlaPheLeuAspCysGluProIleuLeu	159
Db	421	GTGGCGCGCCCGCTCGCGGAGGACATTAATGACTTCCTGTCAGTGGATCGATCCCTCTTC	480
Qy	160	ArgPheAspTyrTyrPheProGlnValProGluHisArgSerAlaGluGlnProIleuArg	179
Db	481	CGCTACGATCTTCCCGAAGTGCAGAGACCGCGTCCGCGAAGAAACCTTCCGC	540
Qy	180	MetAlaProHisTyrTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsn	199
Db	541	ATGGAGCCCACTACGACTTACGACCATACGCTCGTGCAACGACACGCTGCGCCAC	600
Qy	200	GlyPheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgPro	219
Db	601	GCTTCGTGAGCTGTGACGTGCGAGGTGACGGAATTCGTGACCTCCGACGCTCCCC	660
Qy	220	AspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValIlys	239
Db	661	GCGCGCATAGTGTCTTCTGCGGCGCGGTGCGCACCTTGCGCACGCGGCGGCAAGGTCAAG	720
Qy	240	AlaProAlaGlyHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThr	259
Db	721	GCGCCCAAGCACCGGCTCAAGCTTCCCGGCGCGCACGCGCGTGGAGACGCGCACG	780
Qy	260	SerSerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProIleuAlaArg	279
Db	781	TGAGACGCTTCTTCTCTGCGCGCGGACCGCACCTTCACTTCAAGCTCAGACAGTCGAGG	840
Qy	280	GluCysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGly	299
Db	841	GAGTGGGGTTTCAGAGTCGCGCATCCCGTGGGACGACAGCGTTCAGGAGATGTGCTTGCC	900
Qy	300	GlyAsnTyrValAsnIleArgArgThrSerIysAla	311
Db	901	GGGAACTATGTACATCGAGGAGATTAACCGCGCG	936
RESULT 11			
AAAF60258	ID	AAAF60258 standard; DNA; 996 BP.	
XX	XX	AAAF60258;	
AC	XX		
DT	XX	11-SEP-2003 (revised)	
DT	XX	27-APR-2001 (first entry)	
DE	XX	Cephalosporium acremonium DACs/DAOCs coding sequence.	
XX	XX		
KM	XX	Cephalosporium acremonium; DACs/DAOCs; expandase/hydroxylase;	
KM	XX	Cephalosporium deaceoxycephalosporin C; synthetase/hydroxylase;	
KW	XX	cephalosporin C; penicillin production; antibiotic production; ds.	


```

XX  Acromonium chrysogenum.
XX
XX  US6180361-B1.
XX
XX  30-JAN-2001.
XX
XX  12-DEC-1988; 88US-00283429.
XX
XX  04-MAR-1987; 87US-00021836.
XX
XX  (ELIL ) LILLY & CO ELI.
XX
XX  Ingolia TD, Queener SM, Samson SM, Skatrud PL.
XX
XX  MPI; 2001-181736/18.
XX
XX  P-PsDB; AAB69572.
XX
XX  Novel recombinant DNA sequence that encodes Cephalosporium
XX  deacetoxycephalosporin C synthetase/hydroxylase polypeptide useful for
XX  construction of expression vectors for improving efficiency and yield of
XX  fermentation.
XX
XX  Claim 2; Col 59-60; 67pp; English.
XX
XX  The present sequence encodes the DACS/DACOS (expandase/hydroxylase)
XX  polypeptide of Cephalosporium acromonium. The invention provides a novel
XX  recombinant DNA sequence that encodes a Cephalosporium
XX  deacetoxycephalosporin C synthetase/hydroxylase polypeptide. The DNA
XX  sequence is useful for production of cephalosporin in a Penicillium host
XX  cell. It is useful for the construction of expression vectors for
XX  improving the efficiency and yield of fermentation involving a wide
XX  variety of penicillin and cephalosporin antibiotic-producing organism.
XX  The expression vector is useful in constructing strains for use by the
XX  pharmaceutical industry and to introduce cephalosporin synthesizing
XX  activities into high-level penicillin producing Penicillium strain. The
XX  DNA sequence is useful for screening genomic libraries of organisms that
XX  produce cephalosporin C or similar compounds for the presence of the
XX  gene. It is useful for preparing labelled probes that are used to find
XX  expandase-encoding DNA sequences in lactam-producing species. (Updated on
XX  11-SEP-2003 to standardise OS field)
XX
XX  SQ Sequence 996 BP; 176 A; 342 C; 313 G; 165 T; 0 U; 0 Other;
XX
XX  Alignment Scores:
XX  Pred. No.: 5.05e-91 Length: 996
XX  Score: 933.50 Matches: 178
XX  Percent Similarity: 70.83% Conservative: 43
XX  Best Local Similarity: 57.05% Mismatches: 90
XX  Query Match: 57.06% Indels: 1
XX  DB: 4 Gaps: 1
XX
10719236-1_232-1164 (1-311) x AAF60258 (1-996)
XX
XX  1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGluIleuHisGlnAsp 20
XX  1 ATGACTTCCAAAGTCCCGCTTCTTCCTCGACGACTTCAAGAGCGGCAAGTCTCCAC 60
XX
XX  21 GluPheArgGlySerLeuArgAspGlyGlyLeuPheThrLeuThrAspCysGlyLeuThr 40
XX  61 GAGCTCGCCGAGCGCTGCACCAAGGATATCTTACTTCAAGAGCGGCGCTGTC 120
XX
XX  41 AspThrGluLeuLeuSerLeuAlaValAspIleValIleAspPheGluHisGlySerGlu 60
XX  121 GACGACGACCACTCGCGCGCTGAGAGCTGCGTGTGACTTTTCAAGAACGAAACGAG 180
XX
XX  61 AlaGluLeuArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
XX  181 GAGGAAAGAGGCGCTGACGCTCGCCGACCGTAAGCCCGCGGCTTCTCTGCGCTC 240
XX
XX  81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
XX  241 GAGTGGAGAGACCGCGCTGCTACCGAGACGCGGCAAGTACTCGGACTACTCGACGTGC 300

```

```

QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluValIleTrpThr 120
DB 301 TACTCCATGAGGATGCGGCGGCACTGTCCTCCGAACCGGCGCTTGAGAGACCTCTGGCAG 360
QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
DB 361 GACTACTTTCGACCGGATGACCGGCGGACCAAGAGATGTCGCGCGCGCTTCTCAACTCT 420
QY 141 ThrGlyThrGluProAspGly--GlyValGluAlaPheLeuAspCysGluProLeuLeu 159
DB 421 GTGGCGCGCCCGCTCGCGGAGAGCATTAATGACTTCGTGAGTGGATCTCCCTCCCTC 480
QY 160 ArgPheArgTyrPheProGlnValProGlnHisArgSerAlaGlnGluGlnProLeuArg 179
DB 481 GCGCTTACGATCTTCCCGAAGTGCAGAGACCGCGTCCGCAAGAGAACCTCTCCGC 540
QY 180 MetAlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsn 199
DB 541 ATGGGACCCCACTACGACTTATGACCATTCACGCTCTGACACGACGACGACGCTGCCAAC 600
QY 200 GlyPheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgPro 219
DB 601 GCGTTCGAGCGCTCAGTGCAGAGGTGAGCGAGATTCGTGACCTCCGACGCTCCCC 660
QY 220 AspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValIys 239
DB 661 GCGGCAAGTGTGTTTGTGGCGCGCGGTGCGCACCTGCGCACCGGCGGCAAGGTCAAG 720
QY 240 AlaProGlnHisValAlaAlaProArgArgArgGlnIleAlaGlySerSerArgThr 259
DB 721 GCGCCCAAGACCGCGGTCAAGTCTCCGGGCGGACGACCGCGTGGAGACGCGGACG 780
QY 260 SerSerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArg 279
DB 781 TCGACGCGCTTCTTCTCGCGCGGACCGCACTTCACCTTCACGTCACGACGTCGAGG 840
QY 280 GlyCysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGly 299
DB 841 GAGTGGGATTTCAACGTCCGATCCCGTGGAGCGCACGAGGTTCAGGAGTGGCTTGGC 900
QY 300 GlyAsnTyrValAsnIleArgArgThrSerIysAla 311
DB 901 CGGAATATGTCAACATCGGAGGATTAAGCCGCGG 936
XX
XX  RESULT 12
XX  AAN81125
XX  ID AAN81125 standard; DNA; 1509 BP.
XX  XX
XX  AAN81125;
XX  AC
XX  XX  24-OCT-2003 (revised)
XX  DT 25-MAR-2003 (revised)
XX  DT 29-OCT-1990 (first entry)
XX  DE Deacetoxycephalosporin-C-synthetase and deacetylcephalosporin-C-
XX  synthetase gene.
XX  KW Deacetoxycephalosporin-C-synthetase; Deacetylcephalosporin-C-synthetase;
XX  cephalosporin; D8 DNA.
XX  OS Acromonium chrysogenum.
XX  PN EP281391-A.
XX  PD 07-SEP-1988.
XX  PF 03-MAR-1986; 88EP-00301845.
XX  PR 04-MAR-1987; 87US-00021836.
XX  PA (ELIL ) LILLY & CO ELI.
XX

```

```
PI Ingolia TD, Queener SW, Samson SM, Skatrud PL;
XX WPI: 1988-251947/36.
DR P-PSDB; AAB80534.
XX
PT Recombinant DNA from Cephalosporium acremonium - encoding
PT deacetylcephalosporin C synthetase - di:acetoxy:cephalosporin C
PT synthetase activity.
XX
PS Disclosure; Page 7; 82pp; English.
XX
CC DNA sequences encoding deacetoxycephalosporin-C-synthetase and
CC deacetylcephalosporin-C-synthetase are used to construct vectors for a
CC wide variety of hosts, eg Escherichia coli, Penicillium and
CC Cephalosporium. Penicillium transformants may be used for cephalosporin
CC production. See also AAB80534, AAN81126, AAN81127, AAN81128 (updated on
CC 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise
CC OS field)
XX
SQ Sequence 1509 BP; 327 A; 456 C; 429 G; 297 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,996-91 Length: 1509
Score: 933.50 Matches: 178
Percent Similarity: 70.83% Conservative: 43
Best Local Similarity: 57.05% Mismatches: 90
Query Match: 57.06% Indels: 1
DB: 1 Gaps: 1

10719236-1_232-1164 (1-311) x AAN81125 (1-1509)

QY 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
DB 400 ATGACTTCCAAAGTCCCGTCTTTCGTCTCGACGACCTCAAGAGCGGCAAGGCTCTCAAC 459
QY 21 GluPheArgArgCysLeuAspAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
DB 460 GAGCTCGCGGAGCGCGTCAACCAAGGGTATCTTCTACTTGACGGAAGAGCGGCTGTGC 519
QY 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGluHisGlySerGlu 60
DB 520 GACGACGACCAACCTCGCGCGGTGACGTCGCTTTTCAGACTTTTTCAGAACGGAAGCGAG 579
QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
DB 580 GAGGAGAGAGAGGCGGTGACGCTCGCCGCAACGCGCGCGGCTTCTCTCCCTC 639
QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
DB 640 GAGTGGAGAGACCGCGCGTGTCAACGAGACGGGCAAGTACTCGGACTACTCGACGCGC 699
QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleThrThr 120
DB 700 TACTCCATGGGATCGCGCGCAACCTGTTCCGAACCGGCGGCTTTCGAGGACGCTCGGAG 759
QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
DB 760 GACTACTTTCGACCGCATGTAACGCGGACCAAGATGTGCGCGCGCGCTTCTCAACTCT 819
QY 141 ThrGlyThrGluProAspGly--GlyValGluAlaPheLeuAspCysGluProLeuLeu 159
DB 820 GTGGGGCGCGCGCTCGCGGAGAGACATGATGACTTCTGTCGAGATGGGATCCCTCTC 879
QY 160 ArgPheArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArg 179
DB 880 CGCCTACGGTACTTCCCGCAAGTGGCGGAGACCGCGTTCGCCGAAGAGAACCCCTCCGC 939
QY 180 MetAlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsn 199
DB 940 ATGGGACCCCACTACGACTATCGACATACGCTGTGTGCACCAAGACCTCCGCCAAC 999
QY 200 GlyPheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgPro 219
```

```
DB 1000 GGCTTCGAGCGCTGACGTGCGAGGTGAGCAAGAAATTCGTGACACTCCGAGCGTCC 1059
QY 220 AspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValLys 239
DB 1060 GCGCCCAATGTCGCTTCTCGCGGCGGTGCGGACCTCTGGCCACGGCGGCAAGTCAAG 1119
QY 240 AlaProGlnHisLeuValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThr 259
DB 1120 GCGCCCAAGCACCGGGTCAAGTCTCCGGCGGACCAAGCGCGGTGCGACAGCCGCACG 1179
QY 260 SerSerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArg 279
DB 1180 TCGAGCGTCTTCTTCTCGCGGCGGACCGCACTTCAAGTTCAGTTCAGTTCAGTTCAGG 1239
QY 280 GluCysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGly 299
DB 1240 GAGGGGGTTTCAACGTCGCCATCCCTCGGACGCGACGAGCGTTCAAGAGAGTGGCTTGGC 1299
QY 300 GlyAsnTyrValAsnIleArgArgThrSerLysAla 311
DB 1300 GGGAACTATGTCAACATCGGAGGATTAAGCCGCGC 1335

RESULT 13
AAF58690
ID AAF58690 standard; DNA; 1511 BP.
XX
AC AAF58690;
XX
DT 11-SEP-2003 (revised)
DT 27-APR-2001 (first entry)
XX
DE Cephalosporium acremonium DACS/DAOS gene.
XX
KW Cephalosporium acremonium; DACS/DAOS; expandase/hydroxylase;
KW Cephalosporium deacetoxycephalosporin C; synthetase/hydroxylase;
KW cephalosporin C; penicillin production; antibiotic production; ds.
XX
OS Acremonium chrysogenum.
XX
PN US6180361-B1.
XX
PD 30-JAN-2001.
XX
PF 12-DEC-1988; 88US-00283429.
XX
PR 04-MAR-1987; 87US-00021836.
XX
PA (EHL ) LILLY & CO ELI.
PI Ingolia TD, Queener SW, Samson SM, Skatrud PL;
DR WPI: 2001-181736/18.
DR P-PSDB; AAB69572.
XX
PT Novel recombinant DNA sequence that encodes Cephalosporium
PT deacetoxycephalosporin C synthetase/hydroxylase polypeptide useful for
PT construction of expression vectors for improving efficiency and yield of
PT fermentation.
XX
PS Disclosure; Col 7-12; 67pp; English.
XX
CC The present sequence is the DACS/DAOS (expandase/hydroxylase) gene of
CC Cephalosporium acremonium. The invention provides a novel recombinant DNA
CC sequence that encodes a Cephalosporium deacetoxycephalosporin C
CC synthetase/hydroxylase polypeptide. The DNA sequence is useful for
CC production of cephalosporin in a Penicillium host cell. It is useful for
CC the construction of expression vectors for improving the efficiency and
CC yield of fermentation involving a wide variety of penicillin and
CC cephalosporin antibiotic-producing organisms. The expression vector is
CC useful in constructing strains for use by the pharmaceutical industry and
CC to introduce cephalosporin synthesizing activities into high-level
CC penicillin producing Penicillium strain. The DNA sequence is useful for
CC screening genomic libraries of organisms that produce cephalosporin C or
```

CC similar compounds for the presence of the gene. It is useful for
 CC preparing labelled probes that are used to find expense-encoding DNA
 CC sequences in lactam-producing species. (Updated on 11-SEP-2003 to
 CC standardise OS field)

XX Sequence 1511 BP; 329 A; 456 C; 429 G; 297 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9,01e-91	Length:	1511
Score:	933.50	Matches:	178
Percent Similarity:	70.83%	Conservative:	43
Best Local Similarity:	57.05%	Mismatches:	90
Query Match:	57.06%	Indels:	1
DB:	4	Gaps:	1

10719236-1_232-1164 (1-311) x AAF58690 (1-1511)

```

QY 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAsp 20
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 ATGACTTCGAAGTCCCGCTTCCTGCTCGACGACTCAAGAGCGGCAAGTCTCACC 460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 21 GluPheArgCysLeuArgAspIlyGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 GAGCTCGCGAGCGCCGTACACCAAGGGATCTTCTACTTGACCGAGAGCGGCTGTC 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 41 AspThrGluLeuSerAlaLysAspIleValIleAspPhePheGluHisGlySerGlu 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 521 GACGACGACCAACCTCGCGCGCTGAGACGTGCGACTTTTTCAGGAACGAGACGG 580
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgGlyPheThrGlyLeu 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 581 GAGGAGAAAGGGCCCTGACCGCGGACCGGACCGCCCGCGCTTCTCTGCTCCTC 640
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 641 GAGTGGAGAGACACCGCGCTGTCACCGAGACGGGCAAGTACTCGGACTACTCGAGTGC 700
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 101 TyrSerMetGlyThrAlaAspLeuPheProSerGlyAspPheGluArgIleTyrPthr 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 701 TACTCATGGGCAATCGCGGCAACCTGTTCCGAAACGGGGGCTTCGAGAGCGTGGCAG 760
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 761 GACTACTTGACCGCATGTACGGCGGACGCAAGAGTGTGGCGCGCCGCTTCTCAACTCT 820
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 141 ThrGlyThrGluProAspGly---GlyValGluAlaPheLeuAspCysGluProLeuLeu 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 GTGGGCGCCCGCTCGCGGAGAGACATTGATGACTTGTGCGAGTGCATCCCTCTC 880
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 ArgPheArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArg 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 CGCTTACGGTACTTCCCGAAGTCCGAGAGACCGGCTGCGAAGAGAACCCCTCCCGC 940
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 180 MetAlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsn 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 ATGGGACCCCACTACGACTATCGACCATCAGCTGTGTCACACGACACCTCCGCAAC 1000
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 200 GlyPheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgPro 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1001 GCGCTTGTAGCGCTGACGTGCGAGGTGAGCGAGAAATGTGTGACCTCCCGACGCTCC 1060
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 AspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLys 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1061 GGGCGCATGTCTTCTTGGCGCGGCGGCGACCTTGGCCACCGGGCGGCAAGGTCAAG 1120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 240 AlaProArgHisIleValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThr 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1121 GGGCCCAACACCGCGGTCAAGTCTCCGGCGGAGACCAACCGGCTGCGACGACGCGCAG 1180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 260 SerSerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArg 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1181 TCGAGCGTCTTCTTCTTGGCGGCGAAGCCGACTTCAAGCTTCAACGTGACGACGAGTGCAG 1240

```

```

QY 280 GluCysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGly 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1241 GAGTGGGGTTTCAAGTCCGCAATCCCTCGAGAGCGCACGACGTTTCAAGGAGTGGCTTGGC 1300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 300 GlyAsnTyrValAsnIleArgArgThrSerIysAla 311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1301 GCGAAGTATGTCACATGCGGAGGATGAAGCGCGG 1336

```

RESULT 14

ADM41691/C
 ADM41691 standard; DNA; 4623 BP.

ADM41691;
 17-JUN-2004 (first entry)

XX Acremonium chrysogenum cephalosporin C biosynthetic cefEF and cefG genes.
 DE Acremonium chrysogenum cephalosporin C biosynthetic cefEF and cefG genes.
 XX Cephalosporin C; antibiotic; cefEF; cefG; gene; ds.

OS Acremonium chrysogenum.

FX Key Location/Qualifiers
 FT CDS complement (122..1118)
 FT /tag= a

FT CDS /gene= "cefEF"
 FT /tag= b
 FT /gene= "cefG"
 FT /note= "Contains introns"
 FT intron /tag= c
 FT intron /tag= 3044..3108
 FT intron /tag= d

PN WO2004026902-A1.

PD 01-APR-2004.

PF 16-SEP-2003; 2003WO-EP010289.

PR 17-SEP-2002; 2002AT-00001397.

PA (SANO) SANDOZ GMBH.

PI Kuernsteiner H, Friedlin E;

DR WPI; 2004-295383/27.

PT Novel Acremonium chrysogenum protein useful in synthetic or semi-
 PT synthetic production of cephalosporin C or its derivatives with
 PT antibiotic properties.

PS Example 6; SEQ ID NO 7; 43bp; English.

XX The present sequence is that of a genomic DNA region of Acremonium
 CC chrysogenum, flanked by XbaI and SmaI BfrI sites, and containing the
 CC cephalosporin C biosynthetic cefEF and cefG genes. The invention relates
 CC to nucleic acids ADM41686-ADM41688 encoding a novel protein ADM41685
 CC involved in cephalosporin C biosynthesis. It is an object of the present
 CC invention to provide a nucleic acid and vectors which code for the new
 CC protein and which can be used for transformation of an A. chrysogenum
 CC host cell such that the host cell is capable of producing cephalosporin C
 CC in good yield. The vector may additionally comprise at least one of the
 CC pcDNA, pCDB, cefD2, cefEF or cefG genes.

XX Sequence 4623 BP; 1040 A; 1284 C; 1319 G; 980 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.25e-90	Length:	4623
Score:	933.50	Matches:	178
Percent Similarity:	70.83%	Conservative:	43
Best Local Similarity:	57.05%	Mismatches:	90

Query Match:	57.06%	Indels:	1
DB:	12	Gaps:	1
10719236-1_232-1164 (1-311) x ADM41691 (1-4623)			
QY	1	MetAepThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp	20
DB	1120	ATGACTTCAGAGTCCCGTCTTGTCTCCAGCACTTCAGAGGCGCAAGGTCTCTACC	1061
QY	21	GluPheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeuThr	40
DB	1060	GAGCTCGCCGAGCGCGTACACCAAGGGTATCTTCTCTGACCGAGAGCGCGCTGCTC	1001
QY	41	AspThrGluLeuLysSerAlaLysAspIleValIleAspPheGlnHisGlySerGlu	60
DB	1000	GACGACGACCACTCGCGCGGTGAGACGTGGTGACTTTTTCAGAGCGAAGGCGAG	941
QY	61	AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu	80
DB	940	GAGGAGAAAGGGCGCTGACCGCTCGACCGGTAAAGCGCGCGCTTCTCTGCCCTC	881
QY	81	GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys	100
DB	880	GAGTGGGAGAGACCGCGCTGTCACCGAGCGGCAAGTACTCGACTCTGACGCGTGC	821
QY	101	TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleThrThr	120
DB	820	TACTCATGGGCATCGCGCGCACTGTCCGAAACCGGGCTTCAGAGAGCTCGGAG	761
QY	121	GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla	140
DB	760	GACTACTTCGACCGCATGACGGCGCAGCAAGATGTGCGCGCGCTTCTCAACTCT	701
QY	141	ThrGlyThrGluProAspGly--GlyValGlnAlaPheLeuAspCysGluProLeuLeu	159
DB	700	GTGGCGCGCCGCTCGCGCGGAGGACATGTGATGCTTCGTCAGTGTGATCCCTCTC	641
QY	160	ArgPheArgTyrPheProGlnValProGlnHisArgSerAlaGluGluGluProLeuArg	179
DB	640	CGCTCAAGTACTTCCCGAAGTGGCGAGGACCGCGTCCCGGAGAGAGACCTTCCGCG	581
QY	180	MetAlaProHisIleTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsn	199
DB	580	ATGGAGCCCACTACGACTATCGACCATCACTCGTGGACCAAGAGCGCTGCGCCAAC	521
QY	200	GlyPheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgPro	219
DB	520	GGCCTTCGAGGCTCGACGTGCGAGGTGACGAGAAATTCGTGACCTCCGACGCTCCC	461
QY	220	AspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLys	239
DB	460	GGCGGCATAGTGTCTTCTCGCGCGCGGTGGCACCTCGGCACGGGCGGCAAGTCAAG	401
QY	240	AlaProArgHisIleValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThr	259
DB	400	GCGCCAGACACCGGGTCAAGTCTCCCGGGCGGACCAAGCGGTGGCAGCAGCGGACG	341
QY	260	SerSerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArg	279
DB	340	TCGACGCTCTTCTTCCTCGCGCGGAGCCGACCTTCAGCTTCACGTCAAGCATCGAGG	281
QY	280	GluCysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGly	299
DB	280	GAGTGGGGTTCAACGTCCGCGCATCCCGTGGAGCGCAGACGACTTCAAGGAGTGGCTG	221
QY	300	GlyAsnTyrValAsnIleArgArgThrSerLysAla	311
DB	220	GGAACATAATGTCAACATGCGGAGGATTAAGCCGGG	185
RESULT 15			
AAQ20267			
ID	AAQ20267 standard; DNA; 956 BP.		
XX			

AC	AAQ20267;
XX	
DT	25-MAR-2003 (revised)
DT	07-APR-1992 (first entry)
XX	
DE	Hydroxylase gene.
XX	
KW	Deacetoxycephalosporin C; DNOC; deacetylcephalosporin C; DAC.
XX	
OS	Streptomyces clavuligerus.
XX	
FH	Key
FT	CDS
FT	1..956
FT	/tag= a
FT	/note= "hydroxylase gene"
XX	
PN	EP465189-A.
XX	
PD	08-JAN-1992.
XX	
PE	01-JUL-1991; 91EP-00305939.
XX	
PR	06-JUL-1990; 90US-00549502.
XX	
PA	(ELIL) LILLY & CO ELI.
XX	
PI	Kovacevic S, Miller JR;
XX	
DR	WP1; 1992-010373/02.
DR	P-PSDB; AAR20110, AAR21420.
PT	DNA encoding Streptomyces clavuligerus hydroxylase - used for increasing
XX	the yield of cephalosporin cpds. and producing new cephalosporin(s).
XX	
PS	Claim 3; Page 4; 15pp; English.
XX	
CC	The sequence encodes a hydroxylase from Streptomyces clavuligerus which
CC	catalyses the reaction in which deacetoxycephalosporin C (DNOC) is
CC	hydroxylated at the 3-methyl group to form deacetylcephalosporin C (DAC).
CC	The cloned hydroxylase gene is useful for increasing the yield of
CC	cephalosporins in fungi and bacteria and to generate new cephalosporins.
CC	The DNA can also be used as a probe to find hydroxylase, expandase, or
CC	expandase/hydroxylase genes in other microorganisms. This sequence
CC	appears to be missing a nucleotide (T) at position 307 resulting in a
CC	shift of reading frame and a completely different protein to that given
XX	in the specification. (Updated on 25-MAR-2003 to correct PA field.)
XX	
SQ	Sequence 956 BP; 156 A; 343 C; 322 G; 135 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	1.56e-89
Score:	919.50
Percent Similarity:	72.64%
Best Local Similarity:	58.63%
Query Match:	56.20%
DB:	2
10719236-1_232-1164 (1-311) x AAQ20267 (1-956)	
QY	2 AapThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGlu 21
DB	7 GACACGCGCCGATACGATCTTCAACCTCGCGCACTGGGAGAGCGCGCATCGAGAGAG 66
QY	22 PheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeuThrAsp 41
DB	67 TTCGCGAGTGGCGACCGGAGTGGGCTTCTTCACTCAACGGGTACGGCGCGGGGAT 126
QY	42 ThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGlnHisGlySerGluAla 61
DB	127 AAGGACCAACCGGCTGGCGACGACGCGCATGTGACTTTCGGAACGACGAGGCC 166
QY	62 GluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlu 81

```

Db      187 GAGAAAGCGCGCTGACACGACGTCCGACCATGCGCGGCTACTCCGCGTGGAG 246
QY      82  SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyr 101
Db      247 GCCGAGAGACCGCCGACGAGTACAGACCGGTTCTTACACGACATCACTGATGTCCTTC 306
QY      102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTPThrGln 121
Db      307 CC-ATGGGATCTCGGGCAACGCTTCCCTCCCGGAGTTCGAGCGGCTGACGAG 365
QY      122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
Db      366 TACTTCGACAGCTCAAGCGCGGCCGACGAGACGCGGCTGCTGACCGCGAGC 425
QY      142 Gly-----ThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeu 158
Db      426 GCGGCTATAGCGCGGATCGTGGAACTGAGACGCTGTGACGCGACCCCGTGG 485
QY      159 LeuArgPheArgTyrPheProGlnValProGluHisArgSerAlaGluGluProLeu 178
Db      486 CTGCGGCTGCGTACTTCCCGAGGTGCGGACGACCGGTCCCGACACAGCCGCGC 545
QY      179 ArgMetAlaProHisTyrAspLeuSerMetValThrLeuIleGlnIleThrProCysAla 198
Db      546 CGGATGCCCCGACACTACGACCTGTGATCATCACTTCATCCACAGACGCGTGC 605
QY      199 AsnGlyPheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArg 218
Db      606 AACGGCTGTGTCAGCTCCAGCGCGGATCGGCGGAACTGTGAGCTGCGCGTGG 665
QY      219 ProAspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyVal 238
Db      666 GAGGACGCGCTGCTGTGATGTGCGGCGGATGCGCCGCTGCGCACCGAGCGCGT 725
QY      239 LysAlaProArgHisIleValAlaAlaProArgArgAspGlnIleAlaGlySerSerArg 258
Db      726 CCCGGCGCCCGGACACGCTCCGCTCCCGCGCGGTATGCGGAGGCGAGCGAC 785
QY      259 ThrSerSerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAla 278
Db      786 ACGTCAGAGGCTCTTCTTCTGCGCCACAGACGACTTCTGTTCTGCGTGGCCAGAGCC 845
QY      279 ArgGluCysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIle 298
Db      846 CCGAGCTACGCGCTCGCTGACCTCGACATGAGAGACGCGCACTTCGCGACTGATC 905
QY      299 GlyGlyAsnTyrValAsnIle 305
Db      906 GGCACCACTACGTCAACATG 926

```

Search completed: September 19, 2005, 15:36:01
 Job time : 638 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: September 19, 2005, 15:23:03; Search time 3691 Seconds
(without alignments)
3207.258 Million cell updates/sec

Title: 10719236-1_232-1164
Perfect score: 1636
Sequence: 1 MDTVPTFSIAELQGLHOD.....ATFQDWIGANYNIRTSKA 311

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_spool/h/10719236-1TRANS/runat_19092005_153157_7751/app_query.fasta.1.45
-DB=EST -QFMT=fastap -SUFPIR=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=10719236-1TRANS.@CGN_1_13437 @runat_19092005_153157_7751 -NCPU=6
-ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	239.5	14.6	656	CF066872	CF066872 Ac5396 Am
2	187	11.4	710	CF066454	CF066454 Ac4235 Am
3	159	9.7	732	BM407972	BM407972 EST582299
4	156	9.5	1382	AY104706	AY104706 Zea maye
5	152.5	9.3	1101	CI960479	CI960479 OsIFCC004
6	146	8.9	632	BM166131	BM166131 BM166131
7	141.5	8.6	1044	CI960315	CI960315 OsIFCC004
8	140.5	8.6	1107	CI978343	CI978343 OsIFCC031
9	135.5	8.3	1059	CI963525	CI963525 OsIFCC009

10	135.5	8.3	1205	7	CK167073	CK167073 FGAS05135
11	135	8.3	768	5	BM141113	BM141113 BM141113
12	133.5	8.2	670	5	B0519124	B0519124 EST626539
13	133.5	8.2	928	7	CV289688	CV289688 aof01-2ms
14	133	8.1	773	5	BM072029	BM072029 BM072029
15	133	8.1	778	5	BM117875	BM117875 BM117875
16	133	8.1	779	5	BM023309	BM023309 BM023309
17	133	8.1	832	5	BM139402	BM139402 BM139402
18	132.5	8.1	642	5	B0149619	B0149619 NF107A05F
19	132	8.1	568	9	TA233D09P	TA233D09P T. brucei
20	131	8.0	721	5	B0858125	B0858125 QG89G12.Y
21	130.5	8.0	562	2	BE776101	BE776101 MY-10-H-0
22	130	7.9	558	4	BG521872	BG521872 15-54 Ste
23	130	7.9	882	7	CV288986	CV288986 aof01-16m
24	129.5	7.9	1023	9	CI982538	CI982538 OsIFSC048
25	129	7.9	749	7	CF666703	CF666703 RTCT1_25
26	128.5	7.9	805	7	CF666848	CF666848 RTCT1_26
27	128.5	7.9	818	7	CO367625	CO367625 RTX1_35 B
28	128.5	7.9	922	7	CP514629	CP514629 CABu0005
29	128.5	7.9	1065	9	CI978345	CI978345 OsIFCC031
30	128	7.8	602	5	B0519125	B0519125 EST626540
31	128	7.8	820	6	CB660961	CB660961 OSJNE030D
32	127	7.8	1053	9	CI963527	CI963527 OsIFCC009
33	125.5	7.7	1098	3	CNS09Y7Z	CBX32857 Arabidops
34	125	7.6	821	8	CC407079	CC407079 PUH1P43TB
35	125	7.6	954	9	CI963528	CI963528 OsIFCC009
36	124.5	7.6	718	1	A0611155	A0611155 A0611155
37	124	7.6	874	7	CO164665	CO164665 FLU1_49-B
38	124	7.6	981	8	CC336687	CC336687 OGOAF35TV
39	124	7.6	1438	3	CNS0A5LR	CBX24519 Arabidops
40	123.5	7.5	720	5	B0743319	B0743319 MHE4102 F
41	123.5	7.5	751	7	CF474443	CF474443 RTW2_20
42	123.5	7.5	1223	7	CK167162	CK167162 FGAS05145
43	123	7.5	668	4	B1929386	B1929386 EST549275
44	123	7.5	816	4	B0414419	B0414419 HVSMK000
45	122.5	7.5	713	2	BE430816	BE430816 SUN006.B0

ALIGNMENTS

RESULT 1
CF066872
LOCUS
DEFINITION CF066872 656 bp mRNA linear EST 08-DEC-2003
AC5396 Amphidinium carterae Amphidinium carterae cDNA Ac5396
ACCESSION
VERSION CF066872
KEYWORDS
SOURCE
ORGANISM
Amphidinium carterae
Amphidinium carterae
Eukaryota; Alveolata; Dinophyceae; Gymnodiniales; Gymnodiniaceae;
Amphidinium.
REFERENCE
1 (bases 1 to 656)
Bachvaroff,T.R., Concepcion,G.T., Rogers,C.R., Herman,E.M. and
Delwiche,C.F.
Dinoflagellate expressed sequence tag data indicate massive
transfer of chloroplast genes to the nuclear genome
Protist 155 (1), 65-78 (2004)
JOURNAL
COMMENT Contact: Charles Delwiche
University of Maryland, College Park
H.J. Patterson Hall, College Park, MD 20742, USA
Tel: 301-405-8300
Fax: 301-314-9082
Email: delwiche@umd.edu
Seq primer: CTCGTGCCGATTCG.
Location/Qualifiers
1..656
/organism="Amphidinium carterae"
/mol_type="mRNA"
/strain="CMP 1314"
/db_xref="taxon:2961"
/clone="Ac5396"

FEATURES
source

Db 490 TTGGCAACAATGCCGT

QY 174 GIUGIUGInProLeuArgMetaIAsnGlyPheValSerLeuMetValThrLeuIleGln 193
 DB 550 GGGGGTCCAGTCGCGCGCATGCCCCGATCAGCAGCTTGCGATCATCTGTGTGAC 609
 QY 194 GlnThrProCysAlaAsnGlyPheValSerLeuGln 205
 DB 610 CAGACCCCGTGGCTAT-GGGTTGTGTGCTGACAA 644

RESULT 3
 BM407972 732 bp mRNA linear EST 10-MAR-2003
 LOCUS EST582299 potato roots Solanum tuberosum cDNA clone cPRO32H16 5'
 DEFINITION end, mRNA sequence.
 ACCESSION BM407972 GI:18259602
 VERSION BM407972.1 GI:18259602
 KEYWORDS Solanum tuberosum (potato)
 SOURCE Solanum tuberosum
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterids; Lamiales; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 732)
 AUTHORS van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S., Uteback, T., Chiemingo, A., Bougri, O., Buell, C.R., Romling, C., Tanksley, S. and Baker, B.
 TITLE Generation of ESTs from potato roots
 JOURNAL Unpublished (2001)
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/orders/ Seg primer: 73.

FEATURES
 source Location/Qualifiers
 1..732
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cPRO32H16"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.26e-07 Length: 732
 Score: 159.00 Matches: 63
 Percent Similarity: 41.02% Conservatave: 42
 Best Local Similarity: 24.61% Mismatches: 111
 Query Match: 9.72% Indels: 40
 DB: 4 Gaps: 11

10719236-1_232-1164 (1-311) x BM407972 (1-732)

QY 20 AspGluPheArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGly--- 38
 DB 38 GATACAAATTAGTAAAGCCTGTTGGAGATAGTGTTCCTCAATTGTAATCATGCTGTA 97
 QY 39 -----LeuThrAspThrGluLeuLysSerAlaIysAspIleValIleAspPhe 55
 DB 98 CCTTTGATCTCAGTCAAGCTCTTAAAGGTCCAG-----TCTTTCTTT 145

QY 56 GluHisGlySerGluAlaGluLysArgAlaValThrSerProValProThr-----Met 73
 DB 146 GAATCTCCAGACAGAGGCAACCTCAATGCT-----TCTCAGCTGCTTAATGTAAGTCTT 199
 QY 74 ArgArgGlyPheThrGlyLeuGluSerGluSerThrAlaGlnIleThrAsnThrGlySer 93
 DB 200 CCTCAGAGCTATACAGAGAGCTTAATCCTTAT-----GAG 238
 QY 94 TyrSerAspTyrSerMetCysTyrSerMetGlyThrAlaAspAsnLeuPheProSerGly 113
 DB 239 TTTAGTAAATTCCTTAATGCTTCCACCTGGTCTCACTTCATTAATCTTCCAGACAAAT 298
 QY 114 -----AspPheGluArgIleTyrThrGlnTyrPheAspArgGlnTyrThrAlaSerArg 131
 DB 299 CCTCTCAATTCAGAGAGCTGATGAAAGATTATCTCCCATTCATGAGATGGAATG 358
 QY 132 AlaValAlaArgGluValLeuArgAlaThrGlyThrGluProAspGlyGlyValAla 151
 DB 359 GTTGTCGAAAGAAATCTTAAGGAGTGTGGGGCTT---CCTCCGTGTGTTGGAAGA 415
 QY 152 Phe-----LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValPro 168
 DB 416 TTCACCAATGACAGAGCTGGGATTCTCATTCCTTAATTTTACTTGGCG----- 466
 QY 169 GluHisArgSerAlaGluGluGlnProLeuArgMetaIAsnGlyPheValSerLeuGlnAlaGluVal 188
 DB 467 -----GCACAGAGAAAGAAAGATAGATGATCAATCAATCAATTAAGATTTCAGTTGC 517
 QY 189 ValThrLeuIleGlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluVal 208
 DB 518 TTCACAAATTGTTTACAA-----AATGAATGGAGGCTGAGAGTTCAAAAG 565
 QY 209 GlyGlyAlaPheThrAspLeuProTyrArgProAspAlaValLeuValPheCysGlyAla 228
 DB 566 GACGCGCAATGAGATCTTAATGAGCCCAACCAATATCCCTGTGTTCACATTTGGAT 625
 QY 229 IleAlaThrLeuValThrGlyGlyGlnValLysAlaProArgHisValAlaAlaPro 248
 DB 626 GCTTTAGAGTCTCTTACGAATGACAAATTCAGAGTCCAGCCATAGATGAGACCA 685
 QY 249 ArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhePhe 264
 DB 686 AAT-----GGAGATCAAGGAATTCATTGGTCTTC 718

RESULT 4
 AY104706 1382 bp mRNA linear HTC 16-OCT-2002
 LOCUS Zea mays PC0076122 mRNA sequence.
 ACCESSION AY104706
 VERSION AY104706.1 GI:2120784
 KEYWORDS HTC.
 SOURCE Zea mays
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 1382)
 AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
 TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 1382)
 AUTHORS Coe, E.H.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
 COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at WSI, mabemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.
 Location/Qualifiers
 source
 1..1182
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:634204"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/Dupont Cornsensus library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores:
 Pred. No.: 7.14e-07 Length: 1382
 Score: 156.00 Matches: 78
 Percent Similarity: 38.02% Conservative: 49
 Best Local Similarity: 23.35% Mismatches: 119
 Query Match: 9.54% Indels: 88
 DB: Gaps: 16

10719236-1_232-1164 (1-311) x AY104706 (1-1382)

QY 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGluPheArg 23
 Db 201 ACCTCAGCGAGCGTCTCCCTGCGCATCGTCCCTCCCTCCGCGCGCGAGTCCGG 260
 QY 24 ArgCysLeu-----ArgAspLysGlyLeuPheThyLeuThrAspCysGlyLeu 39
 Db 261 CGCCCTCTCTCCGAGCGCGGAGAGATCGCTTCTTCCAGGTGCGAACCGCGGCTC 320
 QY 40 ThrAspThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySer 59
 Db 321 TCCCTGAGGCGAGTACAGAGATGAGACGCGTGTGCCAGAGTCTTCCGG-----CTG 374
 QY 60 GluAlaGluLysAlaValThrSerProValProThrMetArgAlaGlyPheThrGly 79
 Db 375 CCGCGCGAGACAGAGCGC-----GGG 395
 QY 80 LeuGluSerGluSerThrAlaGlnIleThr-----AsnThr 91
 Db 396 CTGACTCTGGAGGACAGCGGAGGCGACCGGACTACTCTCCAGCACCATTGTGACACG 455
 QY 92 GlySerTyrSerAspTyrSerMetCysTyrSerMetGlyThrAlaAspAsnLeuPhePro 111
 Db 456 GCGCGCGAGAGTACTGCGCGGACTGCTCGCTGCGCTGCTCC-----TTCGCC 506
 QY 112 Ser---GlyAspPheGluArgIleTrp-----ThrGlnTyrPhe 123
 Db 507 GCCCTCGCGCAGCAGCGCGCGCTGCGCGCAGCAAGCCCGGAGGCTCCGGAGGTGTG 566
 QY 124 AspArgGlnTyrThrAlaSerArgAlaValAlaIleArgIleValLeuArg----- 139
 Db 567 GAGAGGTTACGGTCCAGACCGCGGCGCTGGGATGAGATCTCTCGCTCTGTGGAG 626
 QY 140 AlaThrGlyThrGluProAsp-----GlyGlyValGluAlaPhe 152
 Db 627 GAGCTCGGCGCTCCGCGCGACTACCTGGAGAGGAGCATCAGCGCGCG----- 674
 QY 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGlnHisArgSer 172
 Db 675 -----GACGTCTCTCTCAAGCTCAACCACTACCGCGCTCCCGGACCCCAAGCGC 725
 QY 173 AlaGluGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192
 Db 726 ACG-----CTCGGCGCTCGCGCGACTGTGACCGCAACCTCTCACTTGTGCTC 773
 QY 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyGlyAlaPhe 212

Db 774 -----CTCCCGAGCATGCTCCCTGCG----- 794
 QY 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223
 Db 795 CTGAGAGTCCCTACAGGGGGCGACTGATCAGGTGAGACCTGTGCGCGCTTCTGTC 854
 QY 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValIleValIleAspProArgHis 243
 Db 855 GTCACTTCGGATCCAGCTGAGTGTGACCAAGCGATCTTGAAGACATCGAGCAC 914
 QY 244 HisValAlaIleProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
 Db 915 CCGGTATGACCAAC-----CTGGAGTGGCGCGGACCAAGGTGCGCACG 959
 QY 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283
 Db 960 TTCATATGCCCAACGACGAGCTGCTCATCGC---CCGCGCGCAGATTCTCTACGCGAC 1016
 QY 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrp 297
 Db 1017 GACACCCCGCGGTGCTACCGACCTCCTCCTGCGCACTTC 1058

RESULT 5
 CL960479 1101 bp DNA linear GSS 21-SEP-2004
 LOCUS O81FCC004567 Oryza sativa Express Library Oryza sativa (indica
 DEFINITION cultivar-group) genomic, genomic survey sequence.
 ACCESSION CL960479
 VERSION CL960479.1 GI:52375663
 KEYWORDS GSS.
 SOURCE Oryza sativa (indica cultivar-group)
 ORGANISM Oryza sativa (indica cultivar-group)
 Buiakoyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriaristidae; Oryzaceae; Oryza.
 1 (bases 1 to 1101)
 Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
 Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
 Wong, G. K. S., Deng, X. W. and Wang, J.
 An analysis of transcriptional regulation of the rice genome and
 its comparison to Arabidopsis
 Unpublished (2004)
 CONTACT: Chen Chen
 Department of Bioinformatic
 Beijing Institute of Genomics
 Chinese Academy of Sciences, Beijing 101300, China
 Tel: 86-10-80481559
 Fax: 86-10-80488676
 Email: chenchen@genomics.org.cn
 Rice genomic sequence.
 Class: exon-trapped.

FEATURES
 source
 1..1101
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:39946"
 /clone_lib="Oryza sativa Express Library"
 /note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores:
 Pred. No.: 1.31e-06 Length: 1101
 Score: 152.50 Matches: 70
 Percent Similarity: 38.17% Conservative: 51
 Best Local Similarity: 22.08% Mismatches: 159
 Query Match: 9.32% Indels: 37
 DB: Gaps: 10

10719236-1_232-1164 (1-311) x CL960479 (1-1101)

QY 5 ValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGln-----AspGluPhe 22
 Db 190 ATTCCGGTATCGACTGCGCGAGCTGCGCGGCGAGTGAAGACGAGCTCGAACAACCTC 249

```

QY 23 ArgArgCysLeuArgAspLysGlyLeuPheThyLeuThrAspCysGlyLeuThrAspThr 42
DB 250 AGGCTCGCCCTGGAGACAGTGGGGCTTTCTTCAGGTTGTGACATGAGAGAGAGAG 309
QY 43 GluLeuLysSerAlaLysAspIleValIleAspPhePheGluHisGlySerGluIleVal 62
DB 310 ACGATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
QY 63 LysAlaGlyAlaValThrSerProValProThrMetAlaArgGlyPheThrGlyLeuGluSer 82
DB 370 GAGAAATGACCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 405
QY 83 GluSerThrAlaGlnIleThrAspThrGlySerThrSerAspThrSerMetCysThrSer 102
DB 406 TACGGCCACAGCCTTGTCTTCTTCACAGACCAAGAGAGAGAGAGAGAGAGAGAG 465
QY 103 MetGlyThr-----AlaAspAsnLeuPheProSer-----GlyAsp 114
DB 466 CTCGGCGTCGAGCGCGCGCTTCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 525
QY 115 PheGluValGlyIleThrThrGlyThrPheAspArgGlyIleThrAlaSerAlaValAla 134
DB 526 TTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
QY 135 ArgGluValLeuArgAlaThrGlyThrGlyProAspGlyGlyValGluAlaPheLeuAsp 154
DB 586 GAGCATATCGCGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGAGAGAGAGAGAG 645
QY 155 CysGluProLeuLeuArgPheArgTyrrPheProGlnValProGluHisArgSerAlaGlu 174
DB 646 GCAGTGCAGAGCGCGTGAAGATGAATCTTCAACCGCGCGCGCGCGCGCGCGCG 693
QY 175 GluGlnProLeuArgMetAlaProHisTyrrAspLeuSerMetValThrLeuIleGlnGln 194
DB 694 GAGCTGCTGCTGGCGCTGAGCCCGCATCCGAGCGAGCGCGCGCGCGCGCGCGCG 753
QY 195 ThrProCysAlaAsnGlyPheValSerLeuGlnIleGluValGlyAlaPheThrAsp 214
DB 754 GACGCGCGCGCTCGCGCGCGTGCAGAGTGCCTCGCGCGCGCGCGCGCGCGCG 807
QY 215 LeuProTyrrArgProAspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThr 234
DB 808 GTTCCATCCCTCCCTCCCGCGCGCTCGTCAACGTCGAGAGACCTCGAGAGTGCAC 867
QY 235 GlyGlyGlnValLysAlaProArgHisHis--ValAlaAlaProArgArgAspGlnIle 253
DB 868 AATGGCAGGTACAAAGAGCGTGAAGACCGCGCGCGTGCAGCGAGCGAGACGAC----- 921
QY 254 AlaGlySerSerArgThrSerSerValPhePheLeuArgProAsnAlaAspPheThrPhe 273
DB 922 -----CGCATGTCGTCGTCACCTTCTACGCGCGCGCGCTGACGACTCGAGCTC 969
QY 274 SerValProLeuAlaArgGlyCysGlyPheAspValSerLeuAspGlyGluThrAlaThr 293
DB 970 GGC-----CGGCTGCC-----GAGCTCGTCCGCGAGCGGAGCGCGCGCAG 1011
QY 294 PheGlnAspTrpIleGlyGlyAsnTyrrValAsnIleArgArgThrSerIle 310
DB 1012 TACCGAGCTAACACACGCGAGTACAGCCGCGACTACGTCACGAGCGCG 1062

```

```

REFERENCE 1 (bases 1 to 632)
AUTHORS Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
JOURNAL Unpublished genes in Ciona intestinalis (2002c)
COMMENT Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satou@ascidian.zoo1.kyoto-u.ac.jp.

FEATURES
source
1..632
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rcincol1b07"
/cisue_type="neural complex"
/clone_lib="Nori Satoh unpublished cDNA library, neural complex"

ORIGIN
Alignment Scores:
Pred. No.: 3.36e-06 Length: 632
Score: 146.00 Matches: 38
Percent Similarity: 47.33% Conservative: 24
Best Local Similarity: 29.01% Mismatches: 55
Query Match: 8.92% Indels: 14
DB: 5 Gaps: 3

10719236-1_232-1164 (1-311) x BW166131 (1-632)
QY 159 LeuArgPheArgTyrrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeu 178
DB 444 CTCAATATATCTTTATTTATCCACCAATTAAAGAGAGTGTAAACCCGAACAAA----- 391
QY 179 ArgMetAlaProHisTyrrAspLeuSerMetValThrLeuIleGlnIleThrProCysAla 198
DB 390 GCAGTGGCGGAACATTCAAGATGGGGTTCATCTTGTCTT----- 349
QY 199 AsnGlyPheVal-----SerLeuGlnIleGluValGlyGlyAlaPheThrAsp 214
DB 348 -----TTTGTTCACAACACTGCGCGCTTCAGATTGAGACGGAAGAACTACAAAGAT 295
QY 215 LeuProTyrrArgProAspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValThr 234
DB 294 GTCCCTGTATATGAAGACACCATTTTGTATCAACATTGAGATGCGTGGAAATTATGGACA 235
QY 235 GlyGlyGlnValLysAlaProArgHisHisValAlaAlaProArgArgAspGlnIleAla 254
DB 234 AAAGGAACTAAAGTGCACAAACATCGTATATATATACCCCAAGATGAATGCGGAAGA 175
QY 255 GlySerSerArgThrSerSerValPhePheLeuArgProAsnAlaAspPheThrPheSer 274
DB 174 AATCATCTCGAAGATGATGTATACCTGCTTATCTGTATATGATGTTGTATCAAT 115
QY 275 ValProLeuAlaArgGlyCysGlyPheAspVal 285
DB 114 CAACCTTCGAGTTCAAAGGATGATGACAGATGT 82

```

```

RESULT 6
BW166131/c 632 bp mRNA linear EST 03-NOV-2002
LOCUS BW166131 Nori Satoh unpublished cDNA library, neural complex Ciona
DEFINITION intestinalis cDNA clone rcincol1b07 3', mRNA sequence.
ACCESSION BW166131
VERSION BW166131.1 GI:24523356
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

```

```

RESULT 7
CL960315
LOCUS CL960315 1044 bp DNA linear GSS 21-SEP-2004
DEFINITION OsifCC004224 Oryza sativa Express Library Oryza sativa (Indica
ACCESSION CL960315
VERSION CL960315.1 GI:52375327
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzeae; Oryza.

```

REFERENCE	AUTHORS		TITLE	JOURNAL	COMMENT
(bases 1 to 1044)	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. M., and Wang, J.	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis			
Unpublished (2004)					
Contact: Chen Chen					
Department of Bioinformatic					
Beijing Institute of Genomics					
Chinese Academy of Sciences, Beijing 101300, China					
Tel: 86-10-80481559					
Fax: 86-10-80488676					
Email: chenchen@genomics.org.cn					
Rice genomic sequence.					
Class: exon-trapped.					
Location/Qualifiers					
1..1044					
/organism="Oryza sativa (indica cultivar-group)"					
/mol_type="genomic DNA"					
/db_xref="taxon:39946"					
/clone_lib="Oryza sativa Express Library"					
/note="Oryza sativa exon trapped genomic sequences "					
ORIGIN					
Alignment Scores:					
Pred. No.:	2,34e-05	Length:	1044		
Score:	141.50	Matches:	61		
Percent Similarity:	37.45%	Conservative:	39		
Best Local Similarity:	22.85%	Mismatches:	108		
Query Match:	8.65%	Indels:	59		
DB:	9	Gaps:	10		
10719236-1_232-1164 (1-311) x CL960315 (1-1044)					
Oy	30	GlyLeuPheThyLeuThrAspCysGlyLeuThrAspThrGluLeuYsSerAlaYsAsp	49		
Dd	232	GGGTTCTTCAGGTGTCGGCGACGGCGTCGACGGCGCTTGCGGGCGGGCTCGAC	291		
Oy	50	IleValIleAspPhePheGluHisGlySerGluAlaGluIlyAspAlaValThrSerPro	69		
Dd	292	GGCGCGACGACTTCTTCGGCTCCGGCTCGCGGACGGAAGCGCGCGCG	345		
Oy	70	ValProThrMetArgArgGlyPheThr	78		
Dd	346	GTCCCGCGGACCGGTGTCGGCTACACAGCGCCAGCGCGACTTCGCTCAAGCTC	405		
Oy	79				
Dd	406	CCATGGAAGAGACCTCTTCCTTCGGTTCCACGACCGCGCGCCGCCCTCCGTC	462		
Oy	91	ThrGlySerTyrSerAspTyrSerMetCysTyrSerMetGlyThrAlaAspAsnLeuPhe	110		
Dd	463				
Oy	111	ProSerGlyAspPheGluArgIleTrrPthrGlnTyrPheAspArgGlnTyrThrAlaSer	130		
Dd	502	CCATGCGGG	552		
Oy	131	ArgAlaValAlaArgGluValLeuAlaThrGlyThrGluProAspGlyGlyValGlu	150		
Dd	553	CTGACGATCATGGAACCTCGTGAGCTGAGCTGGGGGTGAGGAGGCTACTATAGGAG	612		
Oy	151	AlaPheLeuAspCysGluProIleuLeuArgPheArgTyrPheProGlnValProGluHis	170		
Dd	613	TTCCTTCGGGACGACGCTCAACATCAGCGGTGACACTACTACCCCGCATCCCGGAG	669		
Oy	171	ArgSerAlaGluGlnGlnProIleuArgMetAlaProHisTyrAspLeuSerMetValThr	190		
Dd	670				
Oy	191	LeuIleGlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGlyGly	210		

Db	721	ATCCTC-----	CTCCAGACGACGCTCGCGGC	747
Oy	211	-----AlapherthAspleuProTyAr	-----ProAspAlaValLeu	223
Db	748	CTCGAGGCTCTCGTCGACGGGGAAATGGGCCCCGTCAGGCCCGCTCCCGGCGCCATGGTC	-----	807
Oy	224	ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLysAlaProArgHis	-----	243
Db	808	ATCAACATCGCGCAGCACCTTCATGCGCGTGTGCAACGGGAGGTATTAAGAGCTGCGAC	-----	867
Oy	244	HisValAlaAlaProArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe	-----	263
Db	868	AGCGCGCTGCTGACCAACGCGCGGAG-----	-----CGCGCTGCTGCGCTTC	912
Oy	264	PheLeuArgProAsnAlaAsp	270	
Db	913	TTCTCTGTCCCGCGGAGGAC	933	
RESULT 8				
LOCUS	CL978343	1107 bp	DNA	linear
DEFINITION	OS:IFCC031672 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.			
ACCESSION	CL978343			
VERSION	CL978343.1	GI:52411187		
KEYWORDS	GS.			
SOURCE	Oryza sativa (indica cultivar-group)			
ORGANISM	Oryza sativa (indica cultivar-group)			
REFERENCE	Eukaryota: Viridiplantae: Streptophyta; Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.			
AUTHORS	1 (bases 1 to 1107) Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.			
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis			
JOURNAL	Unpublished (2004)			
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.			
FEATURES				
source	Location/Qualifiers 1..1107 /organism="Oryza sativa (indica cultivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Express Library" /note="Oryza sativa exon trapped genomic sequences "			
ORIGIN				
Alignment Scores:				
Pred. No.:	3.34e-05	length:	1107	
Score:	140.50	Matches:	69	
Percent Similarity:	41.87%	Conservative:	52	
Best Local Similarity:	23.88%	Mismatches:	117	
Query Match:	8.59%	Indels:	51	
DB:	9	Gaps:	14	
	10719236-1_232-1164 (1-311) x CL978343 (1-1107)			
Oy	5	ValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAspGluPheArg	24	
Db	157	ATTCTTACCGTCGATCTCAGCAAGTTGCTCGATTCATGTCGTCAGACGAGAGACTGA	216	
Oy	25	CysLeu-----ArgAspLysGlyLeuPheTyLeuThrAspCysGlyLeuThr	40	
Db	217	AGCTGGGGCTGCTCGCTTGGAGATGGGGATTCTTTCACTATGATCAACATGGCGTGTG	276	

QY 41 AspThrGluLeuLysSerAlaIleValIleAspPheGluHisGlySerGlu 60
 Db 277 GAGGATGTGATCTCGGACGCGAGGAAGAACAATAGCCGAGTTCTTACGGCTCCAAATGAG 336
 QY 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
 Db 337 ACCAAGAAAG-----GCCGACTCGACGCTGCC-----AGTGGCATC 372
 QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSer---Met 99
 Db 373 GAGGATATAGCCAGCGCTTCGTGGTGTGCGACGACGAGAAAGCTGACCTGACATG 432
 QY 100 CysTyrSerMet-----GlyThrAlaAspAsnLeuPheProSerGlyAspPheGlu 116
 Db 433 TTCTACCTGCTGCTCCGCGCCCGGAGTCACGAAACATG----- 471
 QY 117 ArgIleTyrThrGlnTyr-----PheAspArgGlnTyrThrAla 129
 Db 472 GCCCTCTGGCCCTGCATCTCCATCTTCAGGAATTCGATGATCGGTACTCGCGAG 531
 QY 130 SerArgAlaValAlaArgGluValLeu-----ArgAlaThrGlyThrGluPro 145
 Db 532 ACGGCGAGGGGTGGCAGATGCTGTGGAGTTTCATGCGAAGACATGCGCGTACGCGC 591
 QY 146 AspGlyGlyValGluValPheLeuAspCysGluProLeuLeuArgPheArgTyrPhePro 165
 Db 592 GGGTGGCTGTGGAAAGGTTCCAGAGCCGCGAGCGGATCATGAAACATCAACCCG 651
 QY 166 GlnValProGluHisArgSerAlaGluGlnProLeuArgMetAlaProHisTyrAsp 185
 Db 652 CCGTCCCGGAG-----GCCGCAAGGTGGTCCGCTCGCCGACACCGAC 699
 QY 186 LeuSerMetValThrLeuIleGlnIleThrProCysAlaAsnGlyPheValSerLeuGln 205
 Db 700 GCCGCTCGTACGCTGCTCTCTGAG-----GTGACGACGTCGCGGCTCGGAG 750
 QY 206 ---AlaGluValGlyValAlaPheThrAspLeuProTyrArgProAsp-----AlaVal 222
 Db 751 ATCAGAGAGCCCGGCGAGGTGGCTCGCGCGCCCTCCGACGACGCGCGCTTC 810
 QY 223 LeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValValLeuAlaProArg 242
 Db 811 ATCGTCACAGTCGCGACATCTCGACATCATGACCAACGCGAATACCGAGCGTGGAG 870
 QY 243 HisHis---ValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
 Db 871 CACAGGCGGTGTGGTGGCGCGGACGAG-----CGCGTGTGCGCG 912
 QY 262 ValPhePheLeuArgProAsnAlaAsp 270
 Db 913 GCGGTGTTCACACGACCGCTGCCAGAC 939
 RESULT 9
 CL963525 1059 bp DNA linear GSS 21-SEP-2004
 LOCUS OsTFC009656 Oryza sativa Expressed Library Oryza sativa (indica
 DEFINITION cultivar-group) genomic, genomic survey sequence.
 ACCESSION CL963525
 VERSION CL963525.1 GI:52381774
 KEYWORDS GSS.
 SOURCE Oryza sativa (indica cultivar-group)
 ORGANISM Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 1059)
 Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
 Zhao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
 Wong, G. K. S., Deng, X. W. and Wang, J.
 TITLE An analysis of transcriptional regulation of the rice genome and
 its comparison to Arabidopsis
 JOURNAL Unpublished (2004)

COMMENT Contact: Chen Chen
 Department of Bioinformatic
 Beijing Institute of Genomics
 Chinese Academy of Sciences, Beijing 101300, China
 Tel: 86-10-80481559
 Fax: 86-10-80488676
 Email: chenchen@genomics.org.cn
 Rice genomic sequence.
 Class: exon-trapped.
 FEATURES
 source
 1. 1059
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="genomic DNA"
 /db_xref="taxon:39946"
 /clone_lib="Oryza sativa Expressed Library"
 /note="Oryza sativa exon trapped genomic sequences"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.00012 Length: 1059
 Score: 135.50 Matches: 81
 Percent Similarity: 37.95% Conservative: 34
 Best Local Similarity: 26.73% Mismatches: 127
 Query Match: 8.28% Indels: 62
 DB: 9 Gaps: 13
 10719236-1_232-1164 (1-311) x CL963525 (1-1059)
 QY 5 ValProThrPheSerLeuAlaGluLeu-----GlnGln 15
 Db 151 ATCCCAACCAATAGACGAGGACGACTGAGCCGCGAGCGGTGACGAGTGTGACGAC 210
 QY 16 GlyLeuHisGlnAspGluPheArgArgCysLeuArgAspGlyLeuPheTyrLeuThr 35
 Db 211 GCGCGCAGAGCGGCGGACCTCGGCTCGGCGCTCCAGTCGTGGAGGCTCTTTCGCGTACC 270
 QY 36 AspCysGlyLeuThrAspThrGluLeuLysSerAlaValAspIleValIleAspPhe 55
 Db 271 GGCACGCGCATGCGGAGCCCTTCTCGACAGATCTCGCGCGAGCGGGAATTCTTC 330
 QY 56 GlnHisGlySerGlnAlaGluLysArgAlaValThr---SerProValProThr----- 72
 Db 331 CACCTCCCGCGGAGGAGAGAGAGATGACAAACGTGTGCGCGCGCGACCGCGAC 390
 QY 73 -----MetArgArgGlyPheThrGlyLeuGluSerGluSerThrAla----- 86
 Db 391 GCGCTCGCGCGCGCGCGGAGGTTCCAGCCCGAGGCTTACGACATCGACCGGTGAC 450
 QY 87 -----GlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys----- 100
 Db 451 ACCGACGACGACATCTCTGACTGTGGACCGGCTTACTTCCAGGTGACCGGAGAG 510
 QY 101 -----TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArg 117
 Db 511 GAGCGCGGCTGAGATTCTGGCCGAGACCGGCGG---CGCTCCGGGACTCTCGAG 567
 QY 117 GlnTyrThrGlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluVal 137
 Db 568 GAG-TACACGCGG-----CGAGCGAGCAGAGGTTTCGCGCGGCT 605
 QY 137 IleuArgAlaThrGlyThrGluProAspGlyValAlaGluPheLeuAsp----- 154
 Db 606 GCTGCGCGCCACGCGGAGGTGCTGGGGTTGGGAGAGAGTTCTTCCGCGCAAGGTGCG 665
 QY 155 -----CysGluProLeuLeuArgPheArgTyrPheProGlnValProGluHisArgSe 172
 Db 666 CGAAGAGTGACGACGTACGCGCGTTCACCTACTACCCGCGCATGCGCGG----- 717
 QY 172 AlaGluGlnIleProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192
 Db 718 ---CCGAGGTGTGTGACGCTCAAGCCCAACGACGACGACGACGACGACGCTTC 773
 QY 192 eGlnIleThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyAlaPhe 212

Db 774 CCTCTGCAGACAGACGCTGGCGGC-----CTCAGCTGCTCAAGACGCGCCCTG 824
Qy 212 eThrAspLeuProTyRArgProAspAlaValIleuValPheCysGly----- 227
Db 825 GCTGACATTCGGGTGCTCACCAACGAGAGCTCTGCTGTCGCCGCCGACGAGATTGAGTT 884
Qy 228 ----AlaIleAlaThrIleuValThrGlyGlyGlnVal---LysAlaProArgHisIleVal 245
Db 885 GTTTCATTCGTTGCTGTTCAGATCATGACCAAGTGTTCATGCGCGCGGTGACAGAGGT 944
Qy 245 IAlaAlaProArgHisArgAspGlnIleAlaGlySerSerArgThrSerSerValPhePheLe 265
Db 945 GGTACACGACGACGACGAGG-----AGGATGTCGGTGGTGAATTTCTA 986
Qy 265 uArgPro 267
Db 987 CCAGCCG 993

RESULT 10
CK167073 1205 bp mRNA linear EST 05-DEC-2003
LOCUS FGAS051352 Triticum aestivum FGAS: Tal7 Triticum aestivum cDNA,
DEFINITION mRNA sequence.
ACCESSION CK167073 GI:39000780
VERSION CK167073.1 GI:39000780
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poidea; Triticeae; Triticum.
1 (bases 1 to 1205)
Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
Peniket, C., Roach, J.L. and Sarhan, F.
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_estecs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [60,850].
Plate: Tal711 row: G column: 16.
location/Qualifiers
1..1205
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="wheat line CI 14106"
/db_xref="taxon:4565"
/lab_host="DH5 alpha"
/clone_lib="Triticum aestivum FGAS: Tal7"
/note="Organ: Crown; Vector: pGEM-T; SSH (suppression
subtractive hybridization) cDNA library from genotype
CI14106 cold hardened at 2 C for 21 days and 49 days
(equal amount of cDNA pooled together before subtraction,
tester) and subtracted against genotype CI14106
non-hardened (20 C) (driver). Nitro-pyrole amino anchored
oligo-dT priming and non-directional cloning."

ORIGIN
Alignment Scores: 0.000145 Length: 1205
Pred. No.: 135.50 Matches: 71
Score:

Percent Similarity: 38.44% Conservative: 52
Best Local Similarity: 22.19% Mismatches: 122
Query Match: 8.28% Indels: 75
DB: 7 Gaps: 14
10719236-1_232-1164 (1-311) x CK167073 (1-1205)

Qy 4 ThrValProThrPheSerLeuAlaGluLeu-----GlnGlnGlyLeuHisGln 19
Db 259 AGCATTCCTGTGATTCAGACTTGGTAGCGCTCTTGATCCCGAGACTATACGCTCGAGGCG 318
Qy 20 AspGluPheArgArgCysLeuArgAspGlyLeuPheTyRLeuThrAspCysGlyLeu 39
Db 319 GCCAGGCTCAGGTCGGCATGTGAGGAGTGGGGCTTCTTCAGGTGAAAAACATGAATTT 378
Qy 40 ThrAspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGlnHisGlySer 59
Db 379 CCTGACACAGTTCTTGAAAAACATGAGAAAGCACTTCGACACTTCTTCAGGCTCCCTC 438
Qy 60 GluAlaGluLysArgAlaValThrSerProValProThrMetArgGlyPheThrGly 79
Db 439 GATGAGAGAGACGATTTGGCCAGTTGCTCGAGAGACTCCAGGCTATATGCGCAAGATT 498
Qy 80 LeuGluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyRAspTyRSerMet 99
Db 499 GTTGAGTCGGAACAT----- 513
Qy 100 CysTyRSerMetGlyThrAlaAspAsnLeuPhe-----ProSerGlyAspPhe 115
Db 514 ---CAAAGCTCAGTCAGTGTGACAGGCTTTTACCTTACAGTACCCAGCCAGCTCAGATCGT 570
Qy 116 Glu---ArgIleThrP-----ThrGlnTyRAspArgGlnTyRThrAlaSerArgAla 132
Db 571 GAATATGACCATGCGCTTACACTCTCTGACGTTAGAGAA-----TCCATTAGTGTG 624
Qy 133 ValAlaArgGluValLeuArgAlaThrGlyThr-----GluProAspGly 147
Db 625 TACTCTCTGAGCTGATGAGAGGAGTGAAGTGGCTGCTCATGCAATCATCGCTGAACACTTA 684
Qy 148 GlyValGluAlaPheLeuAsp-----CysGluProLeuLeuArgPheArgTyRAsp 165
Db 685 GGGGTTGATTCCTCCAGTATACATCATGAGTACAGGATGAGATGACCTTACCTTCCCT 744
Qy 745 GCATGCCCC-----GTGGCTCATGATTAAGTTTGGGATTTCCACACATTCGCGAC 795
Db 166 GluValProGluHisArgSerAlaGluGlnGlnProLeuArgMetAlaProHisTyRAsp 185
Qy 186 LeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPheValSerLeuGln 205
Db 796 ATTTCTATGCTACGCTGATGAGTGAATATATGTTGGAGGCTCCANGATCAGAGA 855
Qy 206 AlaGluValGlyLysAlaPheThrAspLeuProTyRArgProAspAlaValLeuValPhe 225
Db 856 NCAGAT-----GCATGGGTGCGCGGTAAACCCACCCCAAGGCAATGGTGGTAAT 906
Qy 226 CysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValLysAlaProArgHisIleVal 245
Db 907 GTTGGTGACTTCTTGTGAGTAATAGACAAACGAAATCAAAAGCATTTAGACACAGAGTC 966
Qy 246 ---AlaAlaProArgHisArgAspGlnIleAlaGlySerSerArgThrSerSerValPhePhe 264
Db 967 ACCGNTATCTCTCAGAGGAGAGATGATCAATTTCTGCT----- 1005
Qy 265 LeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPheAsp 284
Db 1006 -----TTTCACCTTCCA-----AAGTTTGAAT 1026
Qy 285 ValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGlyAsnTyRValAsn 304
Db 1027 ATGAGCTTG-----GGCCAAATATCTGAAT 1050

RESULT 11
BM141113/c

LOCUS BW141113 768 bp mRNA linear EST 03-NOV-2002
 DEFINITION BW141113 Nori Satoh unpublished cDNA library, gastrula and neurula
 Ciona intestinalis cDNA clone rcign058c02 3', mRNA sequence.
 ACCESSION BW141113
 VERSION BW141113.1 GI:24498338
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.
 REFERENCE 1 (bases 1 to 768)
 Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 TITLE Expressed genes in Ciona intestinalis (2002c)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satcho@sci.kyoto-u.ac.jp.
 FEATURES
 source
 1..768
 Location/Qualifiers
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcign058c02"
 /issue_type="whole body"
 /dev_stage="gastrula and neurula"
 /clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"
 ORIGIN
 Alignment Scores: 8.61e-05 Length: 768
 Pred. No.: 135.00 Matches: 62
 Score: 36.24 Conservative: 42
 Percent Similarity: 21.60 Mismatches: 92
 Best Local Similarity: 8.25 Indels: 92
 Query Match: 5 Gaps: 11
 DB: 10719236-1_232-1164 (1-311) x BW141113 (1-768)
 QY 31 LeuPheTYrLeuThraSPCySGlyLeuThraSPThrGluLeuLYSerAlaLYsAla 50
 Db 744 TTGTTTACCTCGTCAATCTGGAATACACAGCACTTCGTTGACGAGTAATAAGTA 685
 QY 51 ValIleASPhePheGluuHISglYserGluAlaGluLYsArgAlaValThrSerProVal 70
 Db 684 GCTTGGATTTCTTCACAAA---CCTGAAGATGAGAAATTAAAGTTTGTAAGAATGTC 628
 QY 71 ProThrMetArgArgGlyPheThrGlyLeuGluSerGlySerThraLaglnIleThraSn 90
 Db 627 AACCAATTC---GGAATACGACTCTATTGGAAAGAAAG---TTAGAT 586
 QY 91 ThrGlySerTYrSerASPyrSerMetCysTYrSerMetGlyThr----- 105
 Db 585 GATGGAAGCCAGGATTAATGAAGCTTCATTAAACATGAGCTTGCGACGTAATACGAC 526
 QY 106 -----AlaSPasnLeuPheProSerGlyASPheGluArgIleThrPthr 120
 Db 525 CCATGATTAATGAGCCGATGATTAATGATTCAC----- 493
 QY 121 GlnTYrPheASPArgGlnTYrThraLaserArgAlaValAlaArgGluValIleuArgAla 140
 Db 492 -----CAATTTCTTCACAC----- 478
 QY 141 ThrGlyThnGluProASPglYglYValGlnAlaPheLeu---ASPcyGluProLeu--- 158
 Db 477 -----GTGAATCATTTAAGAAAGATTGCAAAATGCTTCT 442
 QY 158 ----- 158

Db 441 CTCAGATTTTGATTTCTTTATCATTTGGGCTAAATTTACAGACCAACCAACTATCA 382
 QY 159 -----LeuArgPheArgTYrPhe 164
 Db 381 GCACTTCATGATCGATTCTATGTGCAAAATATCATCTTCGTTACGCTGCTTATATAT 322
 QY 165 ProGlnValProGluHISArgSerAlaGluGlnGlnProLeuArgMetAlaProHisTYr 184
 Db 321 CCA-----CCTATTACACAGAGTTAGAGAAAGGTCAAAATCGATTGCTTGAACACACC 268
 QY 185 ASPLeuSerMetValThrLeuIleGlnIleThraProCysAlaAsnGlyPheValSerLeu 204
 Db 267 GATTATGGCTTAATGACACTTCTCTTCAAGAT-----TCTGTTGGTGAATGCAATC 214
 QY 205 GlnAlaGluValIleGlyLYAlaPheThraSPLeuProTYrArgProASPAlaValIleuVal 224
 Db 213 AAATCGCA---GATGGACTATATATGATCTTGTCCAAATACAGACACATGTTTGATA 157
 QY 225 PheCySGlyAlaIleAlaThreValI-ThrGlyGlyGlnValIysAlaProArgHisHI 244
 Db 156 AA-TGTGGACAGATTTGTTTACAAATCTGGACACATGAGACCTAAATCTTACGCCACATCG 98
 QY 244 sValAlaAlaProArgArgArgGlnIleAlaGlySerSerArgThrSerSerValPhePh 264
 Db 97 CATCTGGTTCAGATGATATACAACTAAGAGCGCTACCCAGACATGATTCCTTCTT 38
 QY 264 eLeuArgProASPsnAlaasp 270
 Db 37 TGTTCATCTCTTGTGAGAC 19
 RESULT 12
 BOS19124
 LOCUS BOS19124 670 bp mRNA linear EST 07-MAR-2003
 DEFINITION EST626539 Generation of a set of potato cDNA clones for microarray
 analyses mixed potato tissues Solanum tuberosum cDNA clone STMJ007
 5' end, mRNA sequence.
 ACCESSION BOS19124
 VERSION BOS19124.1 GI:21377993
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 670)
 Buehl, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
 Karamcheyva, S.A.
 TITLE Generation of a set of potato cDNA clones for microarray analyses
 JOURNAL Unpublished (2002)
 COMMENT Other ESTs: EST626540
 Contact: Robin Buehl
 The Institute for Genomic Research
 9712 Medical Center Dr. Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics
 Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: T3.
 FEATURES
 source
 1..670
 Location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec or Bintje"
 /db_xref="taxon:4113"
 /clone="STMJ007"
 /issue_type="mixed tissues"
 /lab_host="SOLR"
 /clone_lib="Generation of a set of potato cDNA clones for
 microarray analyses mixed potato tissues"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; supplier: Combination of untreated and Phycoflora
 infestans-treated libraries of stolons, leaves, leaflets,

axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."

ORIGIN

Alignment Scores:
Pred. No.: 0.000106 Length: 670
Score: 133.50 Matches: 55
Percent Similarity: 41.28% Conservatave: 42
Best Local Similarity: 23.40% Mismatches: 103
Query Match: 8.16% Indels: 35
Gaps: 10

10719236-1_232-1164 (1-311) x BQ519124 (1-670)

```

Qy 20 AspgluPheargAgysleuArgAsplyselyleuphetYrleuthraepCysegly--- 38
Db 38 GATACAAATTAGTAAAGCCTGTGGAGATATGGTTTCCAAATTGGAAATGCAATGCTGTA 97
Qy 39 -----LeuthraepThrgluLeuylsSerAlaYsAepIleValIleaspPhePhe 55
Db 98 CCTTTGATCTCATAGTCAAGCTCTAAAGGTGCCAAG-----TCTTTCTTT 145
Qy 56 GluHiseglySerGluAlaGluLysArgAlaValThrSerProValProThr-----Met 73
Db 146 GAATCTCCACAGAACAGCAACTCAATGT-----TCTCCAGCTGCTTACGTACTGTT 199
Qy 74 ArgargglyPheThrglyleugluserGluserThralaGlnIleThraenThrglySer 93
Db 200 CCTGCAGCTTAAACAGAACCTTAATCTTTCTTAT-----GAG 238
Qy 94 TySerAsplyrSerMetCytyrSerMetGlyThralaAepAenLeupheProSergly 113
Db 239 TTTAGTGAATCTCTATTATCTTCCACCTGTCTGACTTCATATCTTCCAGACAAT 298
Qy 114 -----AspPheGluArgIleTrpThrglyThraPheaspArgIlyThralaSerArg 131
Db 299 CCTCTCAATTCCAGAGGTGATGAGAGTATTCTCCCAATTCATGAGATAGATG 358
Qy 132 AlaValAlaArgIleValLeuArgAlaThrglyThrgluProAepglyValGluAla 151
Db 359 GTTGTCAAAAGATCTTAGCGAGTGTGGGGCTT---CTCCCTGTCTTGAAGAA 415
Qy 152 Phe-----LeuaspCysegluProleuLeuArgPheArgIlyrPheProGlnValPro 168
Db 416 TTCAACATGACAGAGCTGGGATTTCTCTCATTTATTATTATCTTGCCG----- 466
Qy 169 GluHisargSerAlaGluGluGlnProleuArgMetAlaProHisTyraSpleuSerMet 188
Db 467 -----GCACACAGAGAAAGAAAGATAGATCAATCCACATTAAGATTTCAGTTGC 517
Qy 189 ValThrleuIleGlnIleThraProCyseAlaHsnGlyPheValSerLeuGlnAlaGluVal 208
Db 518 TTCAAAATGTTTACAA-----AATGAAGTTCGAGCGCTGGAGGTTCAAAA 565
Qy 209 GlyIlyAlaPheThraSpleuProTyraSproAplAlaValleuValPheCyseGlyAla 228
Db 566 GAGCGCGAGATGATCTTAATAGCCCCACACCAAGATGCGCTGTGTGCAACATTGGAT 625
Qy 229 IleAlaThrleuValThrglyGlnVallysalAepOargHis 243
Db 626 GCTCTCAAGTCTTAGCAATGACAAATTCAAGATCCACGCCAT 670

RESULT 13
LOCUS CV289688 928 bp mRNA linear EST 23-SEP-2004
DEFINITION aof01-2ms2-e08 aof01 Asparagus officinalis cDNA clone
ACCESSION CV289688
VERSION CV289688.1 GI:52574247
KEYWORDS EST.
SOURCE Asparagus officinalis (garden asparagus)
ORGANISM Asparagus officinalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

REFERENCE

dePamphilis, C., Carlson, J., Ma, H., Soltis, D., Soltis, P.,
Oppenheimer, D., Frohlich, M., Doyle, J., Tankley, S., Webb, M.,
Leebens-Mack, J., Landherr, L., Iluc, D., and Wall, K.
Generation of ESTs from early male inflorescences of Asparagus
officinalis

JOURNAL

Unpublished (2004)

Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 865 6413
Fax: 814 865 9131

Email: cswd3psu.edu or jh110@psu.edu

The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (http://pgn.cornell.edu)

Plate: aof01-2ms2 row: e column: 08

Seq primer: M13F.

FEATURES

source

Location/Qualifiers
1..928
/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="taxon:4686"
/clone="aof01-2ms2-e08"
/tissue_type="male inflorescences"
/lab_host="SOLR"
/clone_lib="aof01"
/note="Vector: pBluescript SK (+/-); site.1: EcoRI;
site.2: XhoI; this is a directionally cloned,
non-normalized library. This library has been generated by
the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program
(DBI-0115684). More information about the project can be
obtained at http://fgp.bio.psu.edu"

ORIGIN

Alignment Scores:

Pred. No.: 0.00017 Length: 928
Score: 133.50 Matches: 61
Percent Similarity: 38.73% Conservatave: 49
Best Local Similarity: 21.48% Mismatches: 127
Query Match: 8.16% Indels: 47
Gaps: 11

10719236-1_232-1164 (1-311) x CV289688 (1-928)

```

Qy 2 AspThrThraValProThraPheSerLeuAlaGluLeuGlnGlnlyleuHnIleGlnAepGlu 21
Db 157 GACATCCCGTCTTGAACCTCTCTGCGCGAC-----CGCGACAGT 198
Qy 22 PheArgArgCyseleuArgAsplys-----GlyleuPheThraeuthraepCyse 37
Db 199 GTGCTCAAGCGCATACCGACCGCTGATGAGATGGGATATTTCAGTTGGTGAACAC 258
Qy 38 GlyleuThraSpleuThrgluLeuylsSerAlaYsAepIleValIleaspPhePheGluHis 57
Db 259 GGGATCCCGACGAGGTGATAGAGAGCTGACAGAGGTGGGACGAGATTTCTTC 312
Qy 58 GlySerGluAlaGluLysArgAlaValThrSerProValProThraeuthraepGlyPhe 77
Db 313 GGGCTCCCGCAGAGAGAGAGAGGTCTACCGCAGCTGCCGGC---TCGGGAGACTTC 369
Qy 78 ThrglyleuGluSerGluSerThralaGlnIleThraenThrglySerTyraSpleuTyra 97
Db 370 CAGGGGAGCGGACCTAAGCTGACAGAGATTTGAGAGGAGAGAGCTTGGGTGATTAT 429
Qy 98 SerMetCytyrSerMetGlyThralaAepAenLeupheProSerglyaspPheGlu--- 116
Db 430 TTGTTT-----CACACGTGTGCGCGAGAGAGAGGAGGTGATTAAC 468

```


QY 117 ArgIleThrThrGlnIleTyrPheAspArgGlnIleTyrThrAlaSerArgAlaValAlaArgGlu 136
 Db 469 AAGTATTGGCCACGACGACCCCTCCGAGTATAGAAAGGCTAACGAAATATACCAAGAC 528
 QY 137 ValLeuArgAlaThr-----GlyThrGluPro 145
 Db 529 CTCCTTCAGATGTCGATGAATGCTGAGCAGCTTATCGTCGTCGATTAAGACCT 588
 QY 146 AspGlyGlyValGluAlaPhe-----LeuAspCysGluProLeuLeuArgPheArgTyr 163
 Db 559 AACGACTTAAAGATGACGTCCGAGCGCATCTCGAGTACCTGCTGAATCAACTAC 648
 QY 164 PheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAlaProHis 183
 Db 649 TACCGCCGCGCCGACG-----CCCGATCTCGCCCTCGGCGCTTACGCGCAC 696
 QY 184 TyrAspLeuSerMetValThrLeuIleGlnIleTyrProCysAlaAsnGlyPheValSer 203
 Db 697 ACCGACATGTCGGCCATCAATCTC-----GTCCGATGAGTACCGGGG 744
 QY 204 LeuGlnAlaGluValGlyValAlaPheThrAspLeuProTyrArgProAspAlaValLeu 223
 Db 745 CTGACAGTGTTCAGAGATGATCATCTGTTGATGCCAATATATATCCCTAATGCGCTCAT 804
 QY 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValAlaPheArgHis 243
 Db 805 GTTCATATCGATCAAAATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 864
 QY 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
 Db 865 AGAACACGCTCACACAGAG-----AAAGCAGATGTCATGCGCAGTG 909
 QY 264 PheLeuArgPro 267
 Db 910 TTCTGCTCACCA 921

RESULT 14

LOCUS BM072029/c 773 bp mRNA linear EST 20-OCT-2002
 DEFINITION BM072029 Nori Satoh unpublished cDNA library, cleaving embryo Clona intestinalis cDNA clone rcic1097p04 3', mRNA sequence.

ACCESSION BM072029
 VERSION BM072029.1 GI:241173441

KEYWORDS EST.
 SOURCE Clona intestinalis

ORGANISM Clona intestinalis
 Clona intestinalis
 Bukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Clona.

REFERENCE 1 (bases 1 to 773)
 Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 Expressed genes in Clona intestinalis (2002c)
 Unpublished (2002)

JOURNAL Contact: Nori Satoh
 Department of Zoology
 Kyoto University

Authors Sakyo-ku, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113

COMMENT Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
 source location/Qualifiers

1..773
 /organism="Clona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcic1097p04"
 /tissue_type="whole body"
 /dev_stage="cleaving embryo"
 /clone_lib="Nori Satoh unpublished cDNA library, cleaving embryo"

ORIGIN

Alignment Scores:

Pred. No.: 0.000149 Length: 773
 Score: 133.00 Matches: 62
 Percent Similarity: 35.57% Conservative: 44
 Best Local Similarity: 20.81% Mismatches: 101
 Query Match: 8.13% Indels: 92
 DB: 5 Gaps: 11

10719236-1_232-1164 (1-311) x BM072029 (1-773)

QY 20 AspGluPheArgArgCysLeuArgAspIleGlyLeuPheTyrLeuThrAspCysGlyLeu 39
 Db 773 GAGGAAATATGCAATGCTTCAAAATGTTGTTTCTACCTCGTCAATATCGAATA 714
 QY 40 ThrAspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGluHisGlySer 59
 Db 713 CACAGCCTTCGTTGACGAGTAAATTAATGCTTGAATTTCTTACAAA--CCT 657
 QY 60 GluAlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGly 79
 Db 656 GAAGATGGAATTAATTAATTTGTGTAAGATTCACAAATTTC--GATACGACTCT 603
 QY 80 LeuGlnSerGlnSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMet 99
 Db 602 ATTGGAAGGAAAG-----TTAGATGATGGAAGCCAGCTGATTAATTAAGCTT 555
 QY 100 CysTyrSerMetGlyThr-----AlaAspAsnLeu 109
 Db 554 TCATTAAACAGGCTTGTGACACGTATACACCCAGCATTAATGACCGAGATTA 495
 QY 110 PheProSerGlyAspPheGluArgIleThrThrGlnIleTyrPheAspArgGlnTyrThrAla 129
 Db 494 GTTCCA-----CAATTTTCTTCC 477
 QY 130 SerArgAlaValAlaArgGluValLeuArgAlaThrGlyThrGluProAspGlyVal 149
 Db 476 ACC-----GTG 471
 QY 150 GluAlaPheLeu---AspCysGluProLeu----- 158
 Db 470 AAATCATTTATGAAGATGCAAAATGCTTCTCTCAGAAATTTGGATTCTTTATCATTT 411
 QY 158 ----- 158
 Db 410 GGGCTAAATTAACAGACCAACCAACTATACAGACTTCATGATGTCGA 351
 QY 159 -----LeuArgPheArgTyrPheProGlnValProGluHisArgSerAla 173
 Db 350 AATATCATTTCTGTTACGGTGCCTTATATATCA-----CTATTAACACAAGATTA 297
 QY 174 GluGlnGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIleGln 193
 Db 296 GAGAAAGGTCAAAATGATGCTTGAACACCCGATTAATGCGCTTAATGACACTTCTTT 237
 QY 194 GlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGlyValAlaPheThr 213
 Db 236 CAAGAT-----TCTGTTGTCGATGTCAGATCAATGCGCA---GATGGCACTTATATC 186
 QY 214 AspLeuProTyrArgProAspAlaValLeuValPheCysGlyAlaIleAlaThrLeuVal 233
 Db 185 GATGCTTGTCCAAATACCGACACACTGTTTGATTA--TGTGCAATTTGTTTCAATATG 127
 QY 234 -ThrGlyGlyGlnValLysAlaPheArgHisIleValAlaAlaProArgAspArgGlnI 253
 Db 126 GACACATGAGACCTTAATTTACGCCACATCGCATCTTGTTCCAGATGATTAACACTTA 67
 QY 253 AlaGlySerSerArgThrSerSerValPhePheLeuArgProAsnAlaAsp 270
 Db 66 GAGGCTAACCCACGACATGATGCTTTCTTTGTCATCTCTGTAGAC 15

RESULT 15

LOCUS BM117875/c 778 bp mRNA linear EST 24-OCT-2002
 DEFINITION BM117875 Nori Satoh unpublished cDNA library, tailbud embryo Clona

intestinalis cDNA clone rcitdb078c22 3', mRNA sequence.
BM117875
 accession: BM117875.1 GI:24364540
 version: EST.
 keywords: Ciona intestinalis
 source: Ciona intestinalis
 organism: Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 phlebobranchia; Clonidae; Ciona.
 1 (bases 1 to 778)
REFERENCE
 Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 TITLE Expressed genes in Ciona intestinalis (2002c)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel.: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES
 source location/Qualifiers
 1..778
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcitdb078c22"
 /tissue_type="whole animal"
 /dev_stage="tailbud embryo"
 /clone_lib="Nori Satoh unpublished cDNA library, tailbud
 embryo"
ORIGIN
Alignment Scores:
Pred. No.: 0.00015 Length: 778
Score: 133.00 Matches: 62
Percent Similarity: 35.57% Conservative: 44
Best Local Similarity: 20.81% Mismatches: 101
Query Match: 8.13% Indels: 92
DB: 5 Gaps: 11
10719236-1_232-1164 (1-311) x BM117875 (1-778)
QY 20 AspgluPheargArgCysleuargAsplysglyLeuphetYrleuthrAspCysglyLeu 39
DB 777 GAGGAATATATACATGCGTTCAACAATGTTGTTTGTACCTCGTCAATACGGAATA 718
QY 40 ThrspthrGluLeuysSerAlaIysaspIleValIleaspPhepheGluHisGlySer 59
DB 717 CACAGCACTTCGTTGATGACGTAATAATAGTTCGATTTCTTCACAAA--CCT 661
QY 60 GluIaagluysArgAlaValThrSerProValProthrMetArgGlyPheThrGly 79
DB 660 GAAATGAGAAATTAATTAAGTTGGTGAATGTCACAAATTC-----GATACGACTCT 607
QY 80 leuGluSerGluSerThrAlaGlnIleThrAnthrGlySerYrSerAspYrSerMet 99
DB 606 ATTGAAAGGAAAG-----TTAGATGATGGAAGCCAGGATTTATTAAGCTT 559
QY 100 CysTyrSerMetGlyThr-----AlaaspAnleu 109
DB 558 TCATTTAATGATGGCTTGTGACAGTATACGACCATGATTAATGACGATGAATTA 499
QY 110 PheProSerGlyaspPheGluArgIleTrpThrGlnYrPheaspArgGlnYrThrAla 129
DB 498 GTTCCA-----CAATTTCTTCC 481
QY 130 SerArgAlaValAlaArgGluValleuArgAlaThrGlyThrGluProaspGlyGlyVal 149
DB 480 ACC-----GTG 475
QY 150 GluAlaPheLeu---AspCysGluProLeu----- 158
DB 474 AATATCATTTATGAAGAATGCAAAATGCTTGCTCTCAGAAATTTGATCTTATCATTT 415

QY 158 ----- 158
DB 414 GGGCTAAATTTGACAGCAACAATATACGACCTTCATGATGATTTATGTGA 355
QY 159 -----LeuArgPheArgYrPheProGlnValProGluHisArgSerAla 173
DB 354 AATATCATTTCTTCGTTACGGTGCCTTTATATCA-----CCTATTACAGAAAGTTA 301
QY 174 GluGluGlnProLeuArgMetAlaProHisYrAspLeuSerMetValThrLeuIleGln 193
DB 300 GAGAAAGGTCAATATGATGATCTTGAACACACGATTAATGCTTAATGACACTTCTCTT 241
QY 194 GlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyAlaPheThr 213
DB 240 CAAGAT-----TCTGTGTTGATTTGCATCAAAATCCGCA--CATGGACTTATATC 190
QY 214 AspleuProTYrArgProaspAlaValleuValPheCysGlyAlaIleAlaThrLeuVal 233
DB 189 GATGCCGCGCCCAATACAGACACTGTTTGATATAA-TGTGGCAGATTTGTTACAAATACG 131
QY 234 -ThrGlyGlnValAlaValPheAlaProArgHisHisValAlaAlaProArgArgAspGln 253
DB 130 GACACATGAGACCTAAATCTACGCCACATCCGATCTTGTTCCAGATGATCAACTTA 71
QY 253 eAlaGlySerSerArgThrSerSerValPhePheLeuArgProAsnAlaAsp 270
DB 70 GAGGGCTACCCACGACAAATGATGCTTCTTTGTTCATCTTCTTGAGGAC 19

Search completed: September 19, 2005, 17:59:29
Job time : 3701 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 19, 2005, 15:24:33 : Search time 205 Seconds
(without alignments)
2482.352 Million cell updates/sec

Title: 10719236-1_232-1164

Perfect score: 1636
Sequence: 1 MDTVPTFSLAELOGLHOD.....ATFQDWIGNYNIRTSKA 311

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h
-O=/cgn2_1/USPTO_epool_h/10719236-1TRANS/runat.19092005.153158.7766/app.query.fasta_1.45
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=g2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humand40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=10719236-1TRANS_@CGN_1_1_63_@runat.19092005.153158.7766 -NCPU=6 -ICPU=3
-NO WMAP -LANG=ENGLISH -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1636	100.0	1230	4	US-10-105-319-1
2	1210	74.0	939	1	US-08-592-411-14
3	1210	74.0	939	1	US-08-591-501-15
4	1195.5	73.1	942	1	US-08-592-411-16
5	1195.5	73.1	942	1	US-08-591-501-14
6	967.5	59.1	972	2	PCT-US95-04801-2
7	165	10.1	1250	5	US-08-379-556A-9
8	154	9.4	105	1	US-08-592-411-12
9	154	9.4	105	1	US-08-591-501-12
10	153	9.4	1236	2	US-08-379-556A-7
11	149.5	9.1	999	4	US-09-902-540-7970
12	149.5	9.1	5051	4	US-09-902-540-806

13	144	8.8	102	1	US-08-592-411-13	Sequence 13, Appl
14	144	8.8	102	1	US-08-591-501-13	Sequence 13, Appl
15	143	8.7	1211	2	US-08-379-556A-1	Sequence 1, Appl
16	133.5	8.2	1293	3	US-09-454-034-7	Sequence 7, Appl
17	122.5	7.5	1161	4	US-09-371-307-77	Sequence 77, Appl
18	122.5	7.5	1395	2	US-08-553-367A-1	Sequence 1, Appl
19	122.5	7.5	1395	3	US-09-295-306-1	Sequence 1, Appl
20	122.5	7.5	1395	3	US-09-734-719-1	Sequence 1, Appl
21	121	7.4	1490	2	US-08-553-367A-5	Sequence 5, Appl
22	121	7.4	1490	3	US-09-295-306-5	Sequence 5, Appl
23	121	7.4	1490	3	US-09-734-719-5	Sequence 5, Appl
24	120.5	7.4	1170	4	US-09-248-039A-182	Sequence 224, App
25	119.5	7.3	1056	4	US-09-489-039A-2244	Sequence 13457, A
26	117	7.2	1643	4	US-09-270-767-13457	Sequence 11932, A
27	116	7.1	1033	4	US-09-252-991A-11932	Sequence 3, Appl
28	112.5	6.9	1259	2	US-08-553-367A-3	Sequence 3, Appl
29	112.5	6.9	1259	3	US-09-295-306-3	Sequence 3, Appl
30	112.5	6.9	1259	3	US-09-734-719-3	Sequence 3, Appl
31	110.5	6.8	1270	4	US-09-614-912-39	Sequence 39, Appl
32	108.5	6.6	1173	4	US-09-252-991A-3942	Sequence 3942, Ap
33	108.5	6.6	1734	4	US-09-252-991A-3996	Sequence 3996, Ap
34	107	6.5	890	2	US-08-379-556A-3	Sequence 3, Appl
35	104	6.4	966	4	US-09-614-912-43	Sequence 58, Appl
36	103.5	6.3	1066	4	US-09-614-912-37	Sequence 43, Appl
37	103.5	6.3	1361	4	US-09-614-912-37	Sequence 2673, Ap
38	103.5	6.3	1205	4	US-09-248-796A-2673	Sequence 1, Appl
39	102.5	6.3	1555	4	US-09-614-912-1	Sequence 2428, Ap
40	101.5	6.2	1209	4	US-09-248-796A-2428	Sequence 1, Appl
41	100	6.1	1318	4	US-09-719-108-1	Sequence 13, Appl
42	99	6.1	1320	2	US-08-695-412B-13	Sequence 13, Appl
43	99	6.1	1320	3	US-09-255-154D-13	Sequence 13, Appl
44	99	6.1	1320	4	US-10-213-452A-13	Sequence 1, Appl
45	99	6.1	536165	4	US-09-214-808-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-105-319-1
; Sequence 1, Application US/10105319
; Patent No. 6699699
; GENERAL INFORMATION:
; APPLICANT: Yum-Bor Yang
; APPLICANT: Chia-Ii Wei
; APPLICANT: Yuh-Shing Hsu
; TITLE OF INVENTION: Ying-Chieh Tsai
; TITLE OF INVENTION: Mutated penicillin Expandase and Process for Preparing 7-ADCA Usi
; FILE REFERENCE: 6653-017-999
; CURRENT APPLICATION NUMBER: US/10/105,319
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232) ... (1164)
US-10-105-319-1
Alignment Scores:
Pred. No.: 6,43e-196
Score: 1636.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 4
DB: 4
Gaps: 0
10719236-1_232-1164 (1-311) x US-10-105-319-1 (1-1230)
Qy 1 MetaspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlnGlyLeuHieGlnasp 20

Db 232 ATGACACGACGGTCCCACTTCAGCTTGGCCGACCTCCAGAGGCTCTCAGCAGAC 291
Qy 21 GluPheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
Db 292 GAGTCCGACAGTGTCTGAGGACAAAGGGCTCTTCTATCTGACGACCTGGGCTGTACC 351
Qy 41 AspThrGluLeuLysSerAlaIleAspIleValIleAspPhePheGluHisGlySerGlu 60
Db 352 GACACCGAGCTGAAGTCGCGCAAGACATCGCATCGACTTCTTGACACGACGACGAG 411
Qy 61 AlaGlySerArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db 412 GCGGAGAGCGCGCGCTCACTCCGCCCTCCCAACATCGCCGCGGCTTACCGCGGTG 471
Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db 472 GAGTCGAGAGACACCGCCAGATCACCAATACCGGACCTACTCGACTCTCGATGTGC 531
Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleThrThr 120
Db 532 TACTGATGGGACCGCGGACAACTCTCCGTCGGTGACTTGACCGGATCTGGACCC 591
Qy 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db 592 CAGTACTTCGACCGCGCATACCCGCTCCCGCGGCTCGCCGCGGAGGTCTCTGGGGCG 651
Qy 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db 652 ACCGGGACCGAGCCCGACGCGGGGTCCAGAGGCTTCTCGACTGGACCCCTGCTGGG 711
Qy 161 PheArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMet 180
Db 712 TTCGCTACTTCCCGCAGGTCGCCGAGACCGGACGCGGAGAGACACCCCTGGGAGTG 771
Qy 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db 772 GCGCGGACCTAGCATCTGTGATGTCTACCTCATCCAGACAGACCCCTGGCGCAAGCGC 831
Qy 201 PheValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeuProTyrArgProAsp 220
Db 832 TTCGTCACCTTCCAGCGCGAGGTCCGCGCGGTTCACGAGACCTGCCCTTACCGTCCGAC 891
Qy 221 AlaValLeuValPheCysGlyValAlaIleAlaThrLeuValThrGlyGluValIysAla 240
Db 892 GCCGCTCTGCTTCTTGGCGGCGCATCGGACCTGGTGAACCGGCGGCAAGTCAAGGCC 951
Qy 241 ProArgHisIleValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 952 CCCCGGACCATGTGCGGCGCCCGCGGAGGACAGATAGGGGACGACGCGGACCTCC 1011
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db 1012 AGTGTGTTCTTCTCTCGTCCCAACGCGGACTTCACTTCTCCCTGCGCGGCGCAG 1071
Qy 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGlyGly 300
Db 1072 TCGCGCTTCGATGTAGCTTGAACGCGGACGACCGGCACTTCAGAGATTGATCGGGGCG 1131
Qy 301 AsnTyrValAsnIleArgArgThrSerIysAla 311
Db 1132 AACCTAGTGAACATCCGCGGACATCCAAAGGCA 1164

RESULT 2
US-08-592-411-14
; Sequence 14, Application US/08592411
; Patent No. 5726032
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for the Efficient Production of
; TITLE OF INVENTION: 7-ADCA via 2-(Carboxymethylthio)acetyl-7-ADCA and
; NUMBER OF SEQUENCES: 3-(Carboxymethylthio)propionyl-7-ADCA
; 17

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,411
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 939 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: No. 5726032ardie lactandurans
INDIVIDUAL ISOLATE: ATCC 27382
FEATURE:
NAME/KEY: CDS
LOCATION: 1..939
OTHER INFORMATION: /gene= "cefE"
US-08-592-411-14

Alignment Scores:
Pred. No.: 1,99e-142 Length: 939
Score: 1210.00 Matches: 220
Percent Similarity: 83.87% Conservative: 40
Best Local Similarity: 70.97% Mismatches: 50
Query Match: 73.96% Indels: 0
Gaps: 0

10719236-1_232-1164 (1-311) x US-08-592-411-14 (1-939)

Qy 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGlu 21
Db 7 GACCGGACCGTCCGACCTTGCATCTGGCCGAGCTGCTGAGGCTTGCACCGAGAGAG 66
Qy 22 PheArgArgCysLeuArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeuThrAsp 41
Db 67 TTCGCGCAGTGCCTGGCGGAGAGGGCGGTGTCTTACTTCAAGGACACCGGGCTCGCGAG 126
Qy 42 ThrGluLeuLysSerAlaIleAspIleValIleAspPhePheGluHisGlySerGluAla 61
Db 127 GCGGACCAACGCTCGCGCGGAGATGCGGTGACTTCTTGACACCGGACCGGAGGCC 186
Qy 62 GluIysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlu 81
Db 187 GAGAGAGAGCGGTGATGACCGCCATCCGACCATCCGCGCGGATACCGCGGGCTGGAG 246
Qy 82 SerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyr 101
Db 247 TCCGAGAGCACCGCGGACGATCAGAACACCGGCAAGTACCCGACTTACATCGATGCTGAC 306
Qy 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleThrPheGln 121
Db 307 TCGATGGGACCGCGGACCACTGTTCGCCAGGCGGAGTTCAGAGAGGGGTGGAGGAC 366
Qy 122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
Db 367 TACTTCGCGGAGATGACCGCGCTTCGACAGAGCTCGCGGACGAGTCTGACTCGGTC 426
Qy 142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
Db 427 GCGCGGACCGGAGGTGCGCATGAGACGCTTCTCGACTGCGGAAACCCCTGCTGGCGCTG 486
Qy 162 ArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAla 181
Db 487 CGTACTTCCCGCAGGTCGCCGAGATGCGGTGCGCGAGGACGCGCTGCGGATGGCC 546

QY 182 ProHisTyrAspLeuSerMetValThrLeuIleGlnInThrProCysAlaAsnGlyPhe 201
DB 547 CCGGACTAGACCTCTCGATTCGTCACCTCGATTCACCAACCCCTTGGCCGAAACGGGTTCC 606
QY 202 ValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAla 221
DB 607 GTCAGCTCTCAGATCGAGGTGACGGGTCCTATGTGGACATCCCGCGCCGACGGCGCGG 666
QY 222 ValLeuValPheCysGlyValAlaIleAlaThrLeuValThrGlyGlyGlnValIleAspAlaPro 241
DB 667 GTGCTGCTCTCTGCGCGCGCGGTGGCGAGCTGTGGCCGACGCGCGATCAAGGGGCGCC 726
QY 242 ArgHisHisValAlaAlaIleProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
DB 727 AAGCACACAGTGGCGCGCGCGCGCGCGGACCAAGCGGCTGGCGACAGCCGACCTCCAGC 786
QY 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlyCys 281
DB 787 GTGTTCTTCTCGCGCCCAACGGGGAGCTTCCGCTTCTCGGTGCGCGGCGCAGGGAGTGC 846
QY 282 GlyPheAspValSerLeuAspGlyGlnThrAlaThrPheGlnAspTropIleGlyGlyAsn 301
DB 847 GGGTTTCAGATCGACATCCCGCGCGCGCGCGCGGACCTTCGACGACTGATCGCGCGCAC 906
QY 302 TyrValAsnIleArgArgThrSerIleAspAla 311
DB 907 TACATCAACATCCGGAAGACCGCGCGCGCC 936
RESULT 3
US-08-591-501-15
Sequence 15, Application US/08591501
Patent No. 5795733
GENERAL INFORMATION:
APPLICANT: BOVENBERG, ROELOF ARY LANS
APPLICANT: KOEKMAN, BERTUS PIETER
APPLICANT: HOEKEMA, ANDREAS
APPLICANT: VAN DER LAAN, JAN METSKE
APPLICANT: VERWEIJ, JAN
APPLICANT: DE VROOM, ERIK
TITLE OF INVENTION: PROCESS FOR THE EFFICIENT PRODUCTION OF
TITLE OF INVENTION: 7-ADCA VIA 3- (CARBOXYETHYLTHIO) PROPIONYL-7-ADCA
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,501
FILING DATE: 13-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 24615-20065.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSPBRSWSH
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 939 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-591-501-15

Alignment Scores:
Prod. No.: 1,996-142 Length: 939
Score: 1210.00 Matches: 220
Percent Similarity: 83.87% Conservative: 40
Best Local Similarity: 70.97% Mismatches: 50
Query Match: 73.96% Indels: 0
DB: 1 Gaps: 0
10719236-1_232-1164 (1-311) x US-08-591-501-15 (1-939)
QY 2 AspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnIleuHISGlnAspGlu 21
DB 7 GACGGCAGCGGCGGACCTTGATCTGCGCGAGGTGGCGGTTCACACAGAGAGAG 66
QY 22 PheArgArgCysLeuArgAspGlyGlyLeuPheThrIleuThrAspCysGlyLeuThrAsp 41
DB 67 TTCCGCGCACTGCTCGCGGAGAGAGGCGCTGTCTTCACTCAAGGCGACCGGGCTCGCGAG 126
QY 42 ThrGluLeuYsSerAlaIleAspIleValIleAspPhePheGlnHISGlySerGluAla 61
DB 127 GCGGACACGCGCTGCGCGCGGAGATCGCGGTGACCTTTCGACACGAGCGACGAGGCC 186
QY 62 GluYsArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlu 81
DB 187 GAGAGAGAGGCGGTGATGACCGCATCCCGACATCCGCGCGGGTACCGCGGCTGGAG 246
QY 82 SerGluSerThrAlaGlnIleThrAsnThrGlySerYsSerAspTyrSerMetCysTyr 101
DB 247 TCCGAGACACCGCGCGAGATCAGACACCGCGCAAGTACCGCATCTCATGTGCTTAC 306
QY 102 SerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGlnuArgIleThrThrIn 121
DB 307 TCGATGGGACCGCGGACACCTGTCCCGCGCGCGAGTTCGAGAGAGCGGTGGAGAGAC 366
QY 122 TyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThr 141
DB 367 TACTTCGCGGGAGTACCGCGGCTTCGCGAGACGTCGCGGCGGAGTCTACCTCGGTC 426
QY 142 GlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPhe 161
DB 427 GCGCGGAAACCGCGAGTCCGCGATGACGCCCTTCTCTCGACTCGAACCCTCTGCGCGTCG 486
QY 162 ArgTyrPheProGlnValProGlnHISArgSerIleGlnGlnInProLeuArgMetAla 181
DB 487 CGTACTTCCCGAGAGTCCCGAGAGATGCGTGGCGAGGAGACGCCCTCGGATGGCC 546
QY 182 ProHisTyrAspLeuSerMetValThrLeuIleGlnInThrProCysAlaAsnGlyPhe 201
DB 547 CCGGACTAGACCTCTCGATTCGTCACCTCGATTCACCAACCCCTTGGCGGAGGGTTC 606
QY 202 ValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAla 221
DB 607 GTCAGCTCTCAGATCGAGGTGACGGGTCCTATGTGGACATCCCGCGCCGACCGCGCGG 666
QY 222 ValLeuValPheCysGlyValAlaIleAlaThrLeuValThrGlyGlyGlnValIleAspAlaPro 241
DB 667 GTGCTGCTCTCTGCGCGCGCGGTGGCGAGCTGTGGCCGACGCGCGATCAAGGGGCGCC 726
QY 242 ArgHisHisValAlaAlaIleProArgArgAspGlnIleAlaGlySerSerArgThrSerSer 261
DB 727 AAGCACACAGTGGCGCGCGCGCGCGGACCAAGCGGCTGGCGACAGCCGACCTCCAGC 786
QY 262 ValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlyCys 281
DB 787 GTGTTCTTCTCGCGCCCAACGGGGAGCTTCCGCTTCTCGGTGCGCGGCGCAGGGAGTGC 846
QY 282 GlyPheAspValSerLeuAspGlyGlnThrAlaThrPheGlnAspTropIleGlyGlyAsn 301
DB 847 GGGTTTCAGATCGACATCCCGCGCGAGACCGCATCTTCAGACGACTGATGCGGGCAC 906
QY 302 TyrValAsnIleArgArgThrSerIleAspAla 311
DB 907 TACATCAACATCCGGAAGACCGCGCGCGCC 936

Db 907 TACATCAACATCCGAGAACCGCCGCC 936

RESULT 4
US-08-592-411-16
Sequence 16, Application US/08592411
Patent No. 5726032
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Process for the Efficient Production of
TITLE OF INVENTION: 7-ADCA via 2-(Carboxymethylthio)acetyl-7-ADCA and
NUMBER OF SEQUENCES: 17
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,411
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 942 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: No. 5726032ardia lactamurans
INDIVIDUAL ISOLATE: LC 411

FEATURE:
NAME/KEY: CDS
LOCATION: 1..942
OTHER INFORMATION: /gene="cefE"
OTHER INFORMATION: /citation= ([1])
PUBLICATION INFORMATION:
AUTHORS: Coque, J J
AUTHORS: Martin, J F
AUTHORS: Liras, P
TITLE: Characterization and expression in
TITLE: Streptomyces lividans of cefD and cefE genes from
TITLE: No. 5726032ardia lactamurans: the organization of the
TITLE: cephamycin gene cluster differs from that in
TITLE: Streptomyces clavuligerus
JOURNAL: Mol. Gen. Genet.
VOLUME: 236
PAGES: 453-458
DATE: 1993
US-08-592-411-16

Alignment Scores:
Pred. No.: 1,35e-140 Length: 942
Score: 1195.50 Matches: 219
Percent Similarity: 83.28% Conservative: 40
Best Local Similarity: 70.42% Mismatches: 51
Query Match: 73.07% Indels: 1
DB: 1 Gaps: 1

10719236-1_232-1164 (1-311) x US-08-592-411-16 (1-942)

Qy 2 AspThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAspGlu 21
Db 7 GACGCGACCGTGGCCACCTTCGATCTGGCCGAGCTGCGTGAAGGCTTCACACAGAGAG 66
Qy 22 PheArgAGCyGleuArgAspGlyGlyLeuPheTyrLeuThrAspCyGlyLeu---Thr 40
Db 67 TTCGCGACCTGCTCGCGAGAGAGCGGTTCTACCTCAAGGACACCGGCGCTGCCGCC 126
Qy 41 AspThrGluLeuSerAlaLysAspIleValIleAspPhePheGluHisGlySerGlu 60

Db 127 GAGCGGACCAACGCTCGGGCGGGAGATCGGGGTGAGACTTTCGACCAACGACCGAG 186
Qy 61 AlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeu 80
Db 187 GCCGAGAGAGAGCGGTGATGAGCCGATCCGACATCCGCGGCGAGTAACGCCGGCTG 246
Qy 81 GluSerGluSerThrIleGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db 247 GAGTCGAGAGACCCGCGCATACGACACCCGCAAGTACACCACTTCTGATCTG 306
Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
Db 307 TACTCGATGGGACCGCGGACCACTTCCGACGCGGACGAGTTCCAGAAAGCGTGGAG 366
Qy 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuAGla 140
Db 367 GACTACTTCCGGGAGTGTACCGGCTTGACAGACGTGCGGGGAGGTGCTGACTCG 426
Qy 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCyGlyProLeuLeuArg 160
Db 427 GTCGGCGGGAACCCGAGGTGCGCATGAGACGCTTCTCGACTGCGAACCCCTGCTGCC 486
Qy 161 PheArgTyrPheProGlnValProGluHisArgSerAlaGluGlnProLeuArgMet 180
Db 487 CTGGGCTACTTCCCGAGGTGCCGAGGATCGCGTGGCGGAGGACGCGCTGCCGATG 546
Qy 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db 547 GCCCGGACATACACCTCTGATCGTACCTTATCACCAGACCCCTTGCGGAACGGG 606
Qy 201 PheValSerLeuGlnAlaGluValAlaGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db 607 TTGCTGACCTGACGATCGAGTGCAGGTGACGGTCTTATGTGACACTCCCGCGCACCGGGC 666
Qy 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValAla 240
Db 667 GCGGTGTGTGTTTGGCGGCGGTGGCGACCTGTGTGGCCGACCGCGCATCAAGGGC 726
Qy 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 727 CCAGAGACACCATGGCGCGCGCGCGGCGGACAAACGGGTGGCGACGCCACCTCC 786
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgIle 280
Db 787 AGCGTGTCTTCTGCGCCCAACGGGACTTCCGCTCTCGGTGCGCGGCGCAAGAG 846
Qy 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300
Db 847 TGGCGGTTGAGCTGACATCCCGCGCGAGACCGCACCTTCGACGACTGATCGCGGC 906
Qy 301 AsnTyrValAsnIleArgArgThrSerIleAla 311
Db 907 AACTACATCAACATCCGAGAACCGCGCGCC 936

RESULT 5
US-08-591-501-14
Sequence 14, Application US/08591501
Patent No. 5795733
GENERAL INFORMATION:
APPLICANT: BOVENBERG, ROELOF ARY LANS
APPLICANT: KOERMAN, BERTUS PIETER
APPLICANT: HOEKEMA, ANDREAS
APPLICANT: VAN DER LAAN, JAN METSKE
APPLICANT: VERMEIJ, JAN
APPLICANT: DE VROOM, ERIK
TITLE OF INVENTION: PROCESS FOR THE EFFICIENT PRODUCTION OF
TITLE OF INVENTION: 7-ADCA VIA 3-(CARBOXYETHYLTHIO) PROPIONYL-7-ADCA
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC

COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,501
FILING DATE: 13-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 24615-20065.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFORSMH
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 942 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PUBLICATION INFORMATION:
AUTHORS: COQUE et al.,
JOURNAL: Mol. Gen. Genet.
VOLUME: 236
PAGES: 453-458
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 14: FROM 1 TO 942
US-08-591-501-14

Alignment Scores:
Pred. No.: 1,356-140 Length: 942
Score: 1195.50 Matches: 219
Percent Similarity: 83.428 Conservative: 40
Best Local Similarity: 70.428 Mismatches: 51
Query Match: 73.078 Indels: 1
DB: 1 Gaps: 1

10719236-1_232-1164 (1-311) x US-08-591-501-14 (1-942)
QY 2 AapTrrThvaIProThrPheserleuAlaGluLeuGlnGlyLeuHISgluAspGlu 21
Db 7 GACGGACCGCTGCCGACCTTCATCTGGCCGAGCTGCGGAGGCGCTTGACACGAGAGG 66
QY 22 PheArGArGcYsLeuArGAspLySgLyLeuPheTyrLeuThraSpCySgLyLeu---Thr 40
Db 67 TTCCCGCACTGCTCGCGGAGAGAGGCGCTTCTTACTTCAAGGACACCGGCGCTGCCGCGC 126
QY 41 AapTrGluLeuLysSerAlaLysApIleValIleAspPhePheGluHISglYserGlu 60
Db 127 GAGGGGACACGCGCTCGCGGAGAGATCGCGTGACTTCTTGACACGAGACCGAG 186
QY 61 AlaGluYsArGAlaValThrSerProValProThrMetArGArGlyPheThrGlyLeu 80
Db 187 GCCGGAAGAAAGCGGATGATGACCGCGATCCGACCATCCGCGGCGGAGACCGCGGCTG 246
QY 81 GluSerGluSerThraIgluIleThrAnThrGlySerTySerAspTySerMetCys 100
Db 247 GAGTCGAGAGACACCGCGAGATCAAGAAACCGCGCACTACACTGATGCG 306
QY 101 TyrSerMetGlyThraIAspAsnLeuPheProSerGlyAspPheGluArgIleTrpThr 120
Db 307 TACTCGATGGGACACCGCGGACAACTGTCCCAACGCGGAGTTGAGAAGGCGTGGAG 366
QY 121 GlnTyrrPheAspArGlnTyrrThraIAserArGAlaValAlaArgIleValLeuArgAla 140
Db 367 GACTCTCTCGCGGATGATCCGCGCTTGCAGAGACGTTCGCGGCGGAGGTGTGACCTCG 426
QY 141 ThrGlyThrGluProAspGlyGlyValAlaIlePheLeuAspCyGluProLeuLeuArg 160

Db 427 GTCCGCGGGAACCGGAGAGTGGATGACGCGCTTCTCGACTGGAAACCCCTGCTGGC 486
QY 161 PheArGlyrrPheProGlnValProGluHISarGserAlaGluGlnProLeuArgMet 180
Db 487 CTGCGCTACTTCCCGGAGAGTGGCGGAGATGCGCGTGGCGGAGACACCGCTGCGGATG 546
QY 181 AlaProHISrTyrrAspLeuSerMetValThrLeuIleGlnGlnThrProCyAlaAsnGly 200
Db 547 GCCCGGACACTACGACTTCATGCTCAACCTGATCCACGAGACCCCTTGGCGGAACGGG 606
QY 201 PheValSerLeuGlnAlaGluValAlaIleValaPheThrAspLeuProTyrrArGProasp 220
Db 607 TTCCGACACCTGCGAGAGTGAAGTGAAGGCGGCTGATGCGACATCCCGCGGACCGGCGC 666
QY 221 AlaValLeuValaPheCySgLyAlaIleAlaThrLeuValaThrGlyGlyValaIleAla 240
Db 667 GCGGTGCTGTGTCTGCGCGCGGATGCGACGCTGTGGCGGACCGCGGCGATCAAGGCG 726
QY 241 ProArGHSrHISValAlaIleProArGArGAspGlnIleAlaGlySerSerArGThrSer 260
Db 727 CCAGACACACGATGCGCGCGCGCGCGGACGAGACGCGGTGGCGACGACGCGACCTCC 786
QY 261 SerValPhePheLeuArGProAsnAlaAspPheThrPheserValProLeuAlaArgGlu 280
Db 787 AGCGGTCTTCTCTGCGCGCCCAACGGGACTTCCGCTTCTGCTGCGCGGCGGAGGAG 846
QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300
Db 847 TGCGGCTTGCAGCTGACATCCGCGGAGACCGGACCTTGCAGCACTGATGCGGCGC 906
QY 301 AsnTyrrValaenIleArGArGThrSerTyAla 311
Db 907 AACTACATCAACATCCGGAAGACCGCGCGC 939

RESULT 6
PCT-US95-04801-2
Sequence 2, Application PC/TUS9504801
GENERAL INFORMATION:
APPLICANT: Martin, Juan F.
APPLICANT: Coque, Juan R.
APPLICANT: Enguita, Francisco J.
APPLICANT: Fuente, Juan L.
APPLICANT: Ilarena, Francisco J.
APPLICANT: Liras, Paloma
TITLE OF INVENTION: DNA ENCODING CEPHARYCIN BIOSYNTHESIS
TITLE OF INVENTION: LATE GENES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: John W. Wallen III
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04801
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Wallen III, John W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19179
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

; LENGTH: 972 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; PCT-US95-04801-2

Alignment Scores:

Pred. No.:	7.58e-112	Length:	972
Score:	967.50	Matches:	186
Percent Similarity:	73.84%	Conservative:	37
Best Local Similarity:	61.59%	Mismatches:	78
Query Match:	59.14%	Indels:	1
DB:	5	Gaps:	1

10719236-1_232-1164 (1-311) x PCT-US95-04801-2 (1-972)

```

Oy 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGluPheArg 23
    |||||
Db 4 ACGGTACCGGCTCTTCAGCATGGCCGAACCTGGCCGACGGCTGGCCGACGAGATTCCGC 63
    |||||

Oy 24 ArgCysLeuArgAspGlyLeuPheTyrLeuThrAspCysGlyLeuThrAspThrGlu 43
    |||||
Db 64 GAGTGGGCCCGC---CGCGGGGTCTTCTACCTCACCGGGGTACGGGCGCACCGAAGAGAC 120
    |||||

Oy 44 LeuysSerAlaIysAspIleValIleAspPhePheGluHisGlySerGluAlaGluIys 63
    |||||
Db 121 CACCGGTGGCCACACCGCGCATGGACTTCTTCCGCCCAAGGACCGCGGACGAGAGAG 180
    |||||

Oy 64 ArgAlaValThrSerProValProThrMetArgArgGlyPheThrGlyLeuGlnSerGlu 83
    |||||
Db 181 CAGGCGCGTACCAAGATCCCGACCATGGCGCGGGGTACTCGGCGCTGAGGCGGAA 240
    |||||

Oy 84 SerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyrSerMet 103
    |||||
Db 241 AGCACCGCCGAGTTCACCAACACCGCGCATACACCGACTCATCTCATCTCATCTCATG 300
    |||||

Oy 104 GlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThrGlnTyrPhe 123
    |||||
Db 301 GGCATCGCGCGCACCTGTTCCGTCGTAAGAGAGTTCGAGTGGTGGACGCACTACTTC 360
    |||||

Oy 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThrGlyThr 143
    |||||
Db 361 GACGACCTGTACCGCGCGCGGACGAGAGACCGGGCGCTGTGTGACCGCGGACACC 420
    |||||

Oy 144 GluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPheArgTyr 163
    |||||
Db 421 TACGACGCGGAGACCTCGACACCTCTGCTCGACTGCGACCGGCTGCTGCGCTGCGTAC 480
    |||||

Oy 164 PheProGlnValProGlnHisArgSerAlaGluGlnGlnProLeuArgMetAlaProHis 183
    |||||
Db 481 TTCGCGAGGTCCCGAGACACCGCGCGCGGAGTACGAGCCAGCGGATGACCCCGGAC 540
    |||||

Oy 184 TyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPheValSer 203
    |||||
Db 541 TACGACCTGTCCATCATCTTACCAACGACCCCGTGGCGCAAGGTTTCTGTCGAC 600
    |||||

Oy 204 LeuGlnAlaGluValGlyValAlaPheThrAspLeuProTyrArgProAspAlaValLeu 223
    |||||
Db 601 CTGCGAGCGCGAAGTGAAGTGAAGTGAAGCTGCGGACGTCGAGAGACCGCGCTGTG 660
    |||||

Oy 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValIleAlaProArgHis 243
    |||||
Db 661 GTGCTGTGGGCGCATTCGCGCTGGTCAACCGGCGCGGCGCGCGCGCCACACAC 720
    |||||

Oy 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
    |||||
Db 721 CACGTGCTCTCCCGGAGCGGACGAGTGCACAGGACGAGACCGCACTCGAGGCTGTTC 780
    |||||

Oy 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283
    |||||
Db 781 TTCCTGCGCCCGCTCGACCGATTTCACCTTCTCGGTGCGCGGACGCGGACGAGAGTACGCGCTC 840
    |||||
  
```

```

Oy 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspThrIleGlyGlyAsnTyrVal 303
    |||||
Db 841 GACGTACGCTTGACATGAGAGAGGACCTTGGGACTGTGAGTGGAGCAACCTACGTC 900
    |||||

Oy 304 AsnIle 305
    |||
Db 901 ACGATG 906
    |||
  
```

RESULT 7

US-08-379-556A-9

Sequence 9, Application US/08379556A

Patent No. 5859329

GENERAL INFORMATION:

APPLICANT: HOLTON, TIMOTHY A.

APPLICANT: KEAM, LISA A.

TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL

TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSER: SCULLY SCOTT MURPHY & PRESSER

STREET: 400 GARDEN CITY PLAZA

CITY: GARDEN CITY

STATE: NEW YORK

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/379,556A

FILING DATE: 22-MAR-1995

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: DIGILIO, FRANK S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9592

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516)742-4343

TELEFAX: (516)742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1250 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 11..1015

US-08-379-556A-9

Alignment Scores:

Pred. No.:	1.49e-10	Length:	1250
Score:	165.00	Matches:	72
Percent Similarity:	40.19%	Conservative:	55
Best Local Similarity:	22.78%	Mismatches:	135
Query Match:	10.09%	Indels:	54
DB:	2	Gaps:	14

10719236-1_232-1164 (1-311) x US-08-379-556A-9 (1-1250)

```

Oy 5 ValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGluPheArg 24
    |||||
Db 137 GTTCCAGTATGATTCATTTACCGAGCGGATTAACGAATCTTGTGCTTGATATCCAA 196
    |||||

Oy 25 CysLeuArgAspGlyLeuPheTyrLeuThrAspCysGlyLeuThrAspThrGluLeu 44
    |||||
Db 197 GCGAGCAAGATTCGGGTATTTTCAAGTGTGAACCATGGATACCAAGTGAATTCATT 256
    |||||
  
```



```

Qy 45 LysSerAlaLysAspIleValIleAspPhePheGluHisGlySerGluIleGluLysArg 64
Db 257 AGCAATTACAAATATGTGAAAAAGTTCTTGACTTCCACCAAGAA--GAAAAAGAA 313
Qy 65 AlaValThrSerProValProThrMetCArgGlyPheThrGlyLeuGluSerGluSer 84
Db 314 GTCATTTGTAAACCA-----GATCGTTACAAAGGTCGTTGAA----- 349
Qy 85 ThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyrSerMetGly 104
Db 350 -----GGGTATGGAAACAAAGCTTCAAAAAGGCTTGACAGCAAGCAAGCA 394
Qy 105 ThrAlaAspAsnLeuPhe-----ProSerGlyAspPheGluArgIleTrp 119
Db 395 TGGGTGGATCATTTGTTCAATATAGTTTGCCACCTTCTTTATTAATCATCAATTTTGG 454
Qy 120 ThrGln-----TyrPheAsp-----ArgGlnTyrThrAlaSerArg--Ala 132
Db 455 CCAAAGAACCCCTCCTTACAGAGATACAAATGAGAGATACACCAAGCTTGATAGGG 514
Qy 133 ValAlaArgGluValIleu-----ArgAlaThrGlyThrGluProAsp----- 146
Db 515 GTGGCAATATAGTTCCTTGGGCTGTTGTCAAAAAGGCTTGAGTGGAGAAAGATGAAGTG 574
Qy 147 -----GlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPheArg 162
Db 575 AAGCAAGCCTTGGGTGGCAAGACTTGATCTAC-----ATGTGAAATATTAAC 622
Qy 163 TyrPheProGluValProGluHisArgSerAlaGluGluGlnProLeuArgMetAlaPro 182
Db 623 TACTACCCACCATGTCCA-----TGCCCGGACCTTGCTTGGGAGGCCCA 670
Qy 183 HisTyrAspLeuSerMetValThrIleuIleGlnIleThrProCysAlaIleGlyPheVal 202
Db 671 CATACCGACATGTCTTCATCACCATATCTT-----GTCCCGAATGAAGTTCAA 718
Qy 203 SerLeuGlnAlaGluValGlyAlaPheThrAspLeuProTyrArgProAspAlaVal 222
Db 719 GGTCTACAAAGCTTTTAAAGATGTCATGATGATGATGCTTACCTTCTATAGCTCTC 778
Qy 223 LeuValPheCysGlyValAlaIleAlaThrLeuValThrGlyGlyValIleValAspProArg 242
Db 779 ATTATTACATTTGGTACACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 838
Qy 243 HisHisValAlaAlaPheProArgArgArgArgIleAlaGlySerSerArgThrSerSerVal 262
Db 839 CACAGATCAACTGTGATTAAGAG-----AAAACAAGATGTCGTGCCA 883
Qy 263 PhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGly 282
Db 884 GCATTTTGGAGCCACCGCCAGAGTTTGAGTTGGT--CCAAATCCAAAGCTCGTCAT 940
Qy 283 PheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIle 298
Db 941 AAAAGCATCCACCAAAATACAAAGCAAGAACTACAAAGACTATATGTC 988

```

```

; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: oligonucleotide 12
US-08-592-411-12

Alignment Scores:
Pred. No.: 7,01e-11 Length: 105
Score: 154.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.41% Indels: 0
DB: 1 Gaps: 0

10719236-1_232-1164 (1-311) x US-08-592-411-12 (1-105)

Qy 283 PheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGlyAsnTyr 302
Db 1 TTGCATGTCTACCTGGAGCGGACGCGCACGTTCCAGATGGATCGGGGGCAACTAC 60
Qy 303 ValAsnIleArgArgThrSerIleAla 311
Db 61 GTGAACATCCCGCGACATCCAGGCA 87

RESULT 9
US-08-591-501-12
; Sequence 12, Application US/08591501
; Patent No. 579573
; GENERAL INFORMATION:
; APPLICANT: BOVENBERG, ROELOF ARY LANS
; APPLICANT: KOERMAN, BERTUS PIETER
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: VAN DER LAAN, JAN METSKE
; APPLICANT: VERMEIJ, JAN
; APPLICANT: DE VROOM, ERIK
; TITLE OF INVENTION: PROCESS FOR THE EFFICIENT PRODUCTION OF
; TITLE OF INVENTION: 7-ADCA VIA 3-(CARBOXYETHYLTHIO) PROPIONYL-7-ADCA
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,501
; FILING DATE: 13-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 24615-20065.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSMWSH
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs

```

2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822 2823 2824 2825

Db 1015 AAGACGAGGTGTTTCCGAGTAC 1038

RESULT 11

US-09-902-540-7970

Sequence 7970, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Wiegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

PRIOR FILING DATE: 2001-07-10

PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 7970

LENGTH: 999

TYPE: DNA

ORGANISM: Myxococcus xanthus

US-09-902-540-7970

Alignment Scores:

Pred. No.:	9,33e-09	Length:	999
Score:	149.50	Matches:	77
Percent Similarity:	32.49%	Conservative:	39
Best Local Similarity:	21.57%	Mismatches:	118
Query Match:	9.14%	Indels:	123
DB:	4	Gaps:	15

10719236-1_232-1164 (1-311) x US-09-902-540-7970 (1-999)

QY 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40

Db 70 GAAATCGACGACGGCTGTGCGACGACGGCTTCTTACGTACGCGGCGACGCGGTGTC 129

QY 41 AspThrGluLeu-----LysSer 46

Db 130 GCGAGGTGCTCGCGCGGTGAGCGGAGACCGCTTCTTGGCGCGCGGTG 189

QY 47 AlaLysAspIleValIle-----AspPheGluHis 57

Db 190 GCGAAGAGGACCATGCCATGTCGATGGCGGTGCGTGGCGAGCGTGTTCACATC 249

QY 58 GlySerGluAlaGluLysArgAlaValIleThrSerProValProThrMetArgArgIle 76

Db 250 GGGGGGGAG-----CTGACCTCGGGCGCGCGGACCGGAGGAGGGGTG 294

QY 77 PheThrGlyLeuGluSerGluSerThrAlaGlnIleThrAsnThrGly----- 92

Db 295 TATCTGGGACGAGACTCGAAGCGAGCATCCGCGCTAAAGCGCGCTGGCGCTGGCAC 354

QY 93 -----SerTyrSerAspTyr----- 97

Db 355 GCGCGCAACCTGTGCGCGCGCGAGGTGCGGAGCTGCGCGCGCGGTGCTCACTACGTG 414

QY 98 -----SerMetCysTyrSerMetGlyThr 105

Db 415 GCGGCTGTGACGCGCGCGCGCATGCGGTGATGAGGAGGTGCGGTGCGGTGCGGCG 471

QY 106 AlaAspAsnLeuPheProSerGlyAspPheGluAlaGlnIleThrThrGlnTyrPheAsp 125

Db 472 -----CTGGACGCGGACTACTTCGCGCGCG 495

QY 126 GlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAlaThrGlyThrGluPro 145

Db 496 CACTACACGCGCGGACCGGACGCGT----- 519

QY 146 AspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPheArgTyrPhePro 165

Db 520 -----CTCTTCGCGCATCTTCATTAACCC 543

QY 166 GlnValProGluHisArgSerAlaGlnGluInProLeuArgMetAlaProHisTyrAsp 185

Db 544 GCGGAGCCCGCAC-----GAGAGGTGACCTGGGCGGTGCGGACGACCGAC 594

QY 186 LeuSerMetValThrLeuIleGlnIleThrProCysAlaAsnGlyPheValSerLeuGln 205

Db 595 TACGGGCTGTACCGCTGTGCGGACGAC-----AACGC-----GGTTCGAG 642

QY 206 AlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAlaValLeuValPhe 225

Db 643 GTGAAGACGCGCGCGGTGGGTGAGTACCGCGCTGCCGCGGACCGCTGGTGGCAAC 702

QY 226 CysGlyAlaIleAlaThrLeuValThrGlyGlyValLysAlaProArgHisIleVal 245

Db 703 ATCGGCGACATGCTGACCGGATGACGCGGCTGTACCGCTCCACCGCGCGGCGGTG 762

QY 246 AlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerValPhePheLeu 265

Db 763 -----AAGACGTGACGCGGCAAGACGAGCTGTCTTCCACTCTTTC 807

QY 266 ArgProAsnAlaAspPheThrPheSerVal-----ProLeuAlaArgGluCysGlyPheAsp 284

Db 808 GACCGG-----GACTTCGCGCGGAGGTCCACCGCTGACCGCGCGCGGTGAC 861

QY 285 Val-----SerLeuAspGlyGluThrAlaThr 293

Db 862 GTAGATGATGACCGCGCGCGGTGGATGGCGGACGTCACGCTTCAGGGGACG 921

QY 294 PheGlnAspTyrIleGlyIleValAsnTyrValAsnIleArgArgThrSerIle 310

Db 922 TACGCGCATCTGCTGTGCGGCAAGGTGTTCAGAGGTGCTTCAGAAAGTCT 972

RESULT 12

US-09-902-540-806

Sequence 806, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Wiegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

PRIOR FILING DATE: 2001-07-10

PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 806

LENGTH: 5051

TYPE: DNA

ORGANISM: Myxococcus xanthus

US-09-902-540-806

Alignment Scores:

Pred. No.:	1.23e-07	Length:	5051
Score:	149.50	Matches:	77
Percent Similarity:	32.49%	Conservative:	39
Best Local Similarity:	21.57%	Mismatches:	118
Query Match:	9.14%	Indels:	123
DB:	4	Gaps:	15

10719236-1_232-1164 (1-311) x US-09-902-540-806 (1-5051)

QY 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40

Db 4093 GAAATCGACGACGGCTGTGCGACGACGGCTTCTTACGTACGAGGCGGACGCGGTGCC 4152

QY 41 AspThrGluLeu-----LysSer 46

Db 4153 GCGGAGGTGCTGCGCGGTGAGCGGAGACCGCTTCTTGGCGTGGCGCGGTG 4212

```

QY 47 AlAlAspIleValIle-----AspPheGluHis 57
Db 4213 GCGAAGAGGCGCATGTCATGCGGCGTGTGCGGAGCGAGCTGTTCCATCTC 4272
QY 58 GlySerGluAlaGluValArgAlaValThrSerProValProThreMetArgGly--- 76
Db 4273 GGGGGGGAG-----CTGACCTCGGGGCGCCCGGACCGGAGAGAGGGGTG 4317
QY 77 PheThrGlyLeuGluSerGluSerThrAlaGlnIleThrAsnThrGly----- 92
Db 4318 TATCTGGGCGACGAGCTCGAAGCGAGCATCCGCGCTGAAGCGCTGCGCTTGAC 4377
QY 93 -----SerTyrSerAspTyr--- 97
Db 4378 GAGCCCAACTGTGCGCCGCGAGGTGCCGAGCTGCCGCGGCGGTCTGACTACGTG 4437
QY 98 -----SerMetCysTyrSerMetGlyThr 105
Db 4438 GCGGCTGCGACGCGCGCGCGCATGCTGATGAGAGGCGATGGCGCTGAGCTGGGC--- 4494
QY 106 AlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThrGlnTyrPheAspArg 125
Db 4495 -----CTGAGCGCGGACTACTCTCCGCGG 4518
QY 126 GlnTyrThrAlaSerArgAlaValAlaArgIleValLeuArgAlaThrGlyThrGluPro 145
Db 4519 CACTACAGCGGCGAGCCGACGCTG----- 4542
QY 146 AspGlyValGluAlaPheLeuAspCysGluProLeuLeuArgPheArgTyrPhePro 165
Db 4543 -----CTCTCCGATCTTCATACCC 4566
QY 166 GlnValProGluHisArgSerAlaGluGluGlnProLeuArgMetAlaProHisTyrAsp 185
Db 4567 GCGAGCGCCCGAC-----GAGGAGGTGAGCTGCGGCGGAGCACACCGCAC 4617
QY 186 LeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPheValSerLeuGln 205
Db 4618 TACGGGCTGCTACGCTGCTGCGCCAGAC-----MACGC-----GGTTGAG 4665
QY 206 AlaGluValGlyIlePheThrAspLeuProTyrArgProAspAlaValLeuValPhe 225
Db 4666 GTAAGACGCGCGCGCTGGGTGAGGTACCGCGCTGCCGCGGACCGCTGTGTGCAAC 4725
QY 226 CysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValIleValProArgHisHisAla 245
Db 4726 ATCGGCGACATGCTGACCGGATGACAGCGGGGTGACCGCTCCACCGCGCGGGGTG 4785
QY 246 AlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhePheLeu 265
Db 4786 -----AAGAAGCTGAGCGGCAAGACAGGCTGCTTCCCATCTTCTTC 4830
QY 266 ArgProAsnAlaAspPheThrPheSerVal---ProLeuAlaArgIleCysGlyPheAsp 284
Db 4831 GACCCG-----GACTTGGCGCGCGAGGTCCACCCCTGCCACGCGGCGGCGCTGAC 4884
QY 285 Val-----SerLeuAspGlyGluThrAlaThr 293
Db 4885 GTAAGATGATGACCGCGCGCGCTGGGATGGCGGCGGCTTCACGCTTCAGGGCAG 4944
QY 294 PheGlnAspTrpIleGlyIleValAsnTyrValaIleIleArgArgThrSerIle 310
Db 4945 TACGCGCATCTGCTGCGCAAGGTTCATCAAGGTGCTCCCTTAAGAAAGTCGT 4995

```

```

NUMBER OF SEQUENCES: 17
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,411
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: oligonucleotide 13
US-08-592-411-13
Alignment Scores:
Pred. No.: 1.22e-09 Length: 102
Score: 144.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.80% Indels: 0
DB: 1 Gaps: 0
10719236-1_232-1164 (1-311) x US-08-592-411-13 (1-102)
QY 284 AspValSerLeuAspGlyGluThrAlaThrPheGluAspTrpIleGlyIleVal 303
Db 1 GATGTGCTGCTGACGCGGAGCGGACCGCATCTTCAGAGATTGATCGGGGCAACTACGTG 60
QY 304 AsnIleArgArgThrSerIle 310
Db 61 AACATCCGCGCATTCACAG 81
RESULT 14
US-08-591-501-13
Sequence 13, Application US/08591501
Patent No. 5795733
GENERAL INFORMATION:
APPLICANT: BOVENBERG, ROELOF ARY LANS
APPLICANT: KOERMAN, BERTUS PIETER
APPLICANT: HOEKEMA, ANDREAS
APPLICANT: VAN DER LAAN, JAN METSKE
APPLICANT: VERWEIJ, JAN
APPLICANT: DE VROOM, ERIK
TITLE OF INVENTION: PROCESS FOR THE EFFICIENT PRODUCTION OF
TITLE OF INVENTION: 7-ADCA VIA 3-(CARBOXYETHYLTHIO) PROPIONYL-7-ADCA
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,501
FILING DATE: 13-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988

```

REFERENCE/DOCKET NUMBER: 24615-20065.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFORSMW
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-591-501-13

Alignment Scores:
Pred. No.: 1,22e-09 Length: 102
Score: 144.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.80% Indels: 0
DB: 1 Gaps: 0

10719236-1_232-1164 (1-311) x US-08-591-501-13 (1-102)

Qy 284 AspValSerLeuAspGlyLeuThrAlaThrPheGlnAspTrpIleGlyGlyAsnTyrVal 303
Db 1 GATGTACGCTGGAGCGGAGACCGCCACGTTCCAGATTGATCGGGGCAACTACGTG 60

Qy 304 AsnIleArgArgThrSerIys 310
Db 61 AACATCCGCCGACATCCAAAG 81

RESULT 15
US-08-379-556A-1
Sequence 1, Application US/08379556A
Patent No. 5859329
GENERAL INFORMATION:
APPLICANT: HOLTON, TIMOTHY A.
APPLICANT: KEAM, LISA A.
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOL
TITLE OF INVENTION: SYNTHASE ENZYMES AND USES THEREFORE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY SCOTT MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/379,556A
FILING DATE: 22-MAR-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, FRANK S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9592
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1211 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

FEATURE:
NAME/KEY: CDS
LOCATION: 59..1101
US-08-379-556A-1

Alignment Scores:

Pred. No.: 8,36e-08 Length: 1211
Score: 143.00 Matches: 59
Percent Similarity: 38.64% Conservative: 43
Best Local Similarity: 22.35% Mismatches: 114
Query Match: 8.74% Indels: 48
DB: 2 Gaps: 9

10719236-1_232-1164 (1-311) x US-08-379-556A-1 (1-1211)

Qy 27 ArgAspIysGlyLeuPheTyrLeuThrAspCysGlyLeuThrAspThrGluLeuIysSer 46
Db 287 AAAGAGTGGGGATATTCCTCAACTGATCAACCATGGCATCTCGATGAGGCTATCGCGAT 346

Qy 47 AlaIysAspIleValIleAspPheGluThrSerGlyAlaGluIleValArgAlaVal 66
Db 347 TTACGAAAGTAGGAAAGAGTCTTGAACATGTTCCACAGAGAGAAAGAGCTGATT 406

Qy 67 Thr---SerProValProThrMetArgArgGlyPhe---ThrGlyLeuGluSerGluSer 84
Db 407 GCCAAGACTCCAGATCAACAGACATTGAAGGCTATGAACTTCTCTGCAGAGAGATG 466

Qy 85 ThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCysTyrSerMetGly 104
Db 467 GAAAGCAAG-----AAAGGT 481

Qy 105 ThrAlaAspAsnLeuPhe-----ProSerGlyAspPheGluArgIleTrp 119
Db 482 TGGGTGATCATTTGTTCCATTAAGATTGGCTCTTGCCGTCAACTATGCTATTGG 541

Qy 120 -----ThrGlnTyrPheAspArgGlnTyrThr 128
Db 542 CCTAAACCTCTTCATACAGGAGCAACAGAAATATGAAAGAGATGCGAGAA 601

Qy 129 AlaSerArgAlaValAlaArgGluValLeuArgAlaThrGlyThrGluProAspGly 148
Db 602 GTTGAGACAGAAATTTTAAAGAGCTGTCTGGGCTTGGGCTTGAAGGCATGAATG 661

Qy 149 ValGluAlaPheLeuAspCysGlu-----ProLeuLeuAlaPheArgTyrPheProGln 166
Db 662 ATAGAGGAGCTGTGTGATGATGATTACTTGTGAAGATCAACTATTATCCACCA 721

Qy 167 ValProGluIleArgSerAlaGluGlnProLeuArgMetAlaProIleTyrAspLeu 186
Db 722 TGCCCAAG-----CCGATTGGCTCTTGTGTGTGGCCCATACGACATG 769

Qy 187 SerMetValThrLeuIleGlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAla 206
Db 770 TCATATATCACCAATCTT-----GTCCAAATGAAGTCAAGAGCTCCAAAGTG 817

Qy 207 GluValGlyGlyAlaPheThrAspLeuProTyrArgProAspAlaValLeuValPheCys 226
Db 818 TTCAAGGATGGCATTGGTATGATGTCMAAGTACATACCAATGCTTAAATGTCATATT 877

Qy 227 GlyAlaIleAlaThrLeuValThrGlyGlyValIleValIleAspProArgHisIleValAla 246
Db 878 GGTACCAAGTTGATGATCTTACGAAATGCGAAATACAGAGTGTATACATAGACAAAG 937

Qy 247 AlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhePheLeuArg 266
Db 938 GTGAACAAG-----GACAAAGCAAGAATGTATGATGCGGTTTCTTGAG 982

Qy 267 ProAsnAlaAsp 270
Db 983 CCCCCGTGAG 994

Job time : 214 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: September 19, 2005, 16:58:00 ; Search time 741 Seconds

(without alignments)
2796.302 Million cell updates/sec

Title: 10719236-1_232-1164

Perfect score: 1636

Sequence: 1 MDTVPTFTSLAEIQQGLHOD.....ATFGDWIGSNVYNIRTSKA 311

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7389322 seqs, 331285599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.model -DEV=X1h
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000
-USER=10719236-1TRANS @CGN 1.1 480 @runat.19092005.153200.7832/app_query.fasta_1.45
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1636	100.0	1230	16	US-10-105-319-1
2	1636	100.0	1230	18	US-10-719-237-1
3	1636	100.0	1230	18	US-10-719-238-1
4	1636	100.0	1230	19	US-10-719-238-1
5	1636	100.0	1230	20	US-10-789-071-1
6	169	10.3	1101	17	US-10-369-493-35926
7	168	10.3	29870	21	US-10-494-495-1
8	156	9.5	1335	18	US-10-425-114-22510
9	156	9.5	1559	20	US-10-425-114-35660
10	155	9.5	1256	18	US-10-425-114-16150
11	154.5	9.4	1281	19	US-10-437-963-43792
12	154	9.4	1346	18	US-10-425-114-1868
13	154	9.4	1346	18	US-10-425-114-20837
14	154	9.4	1376	18	US-10-425-114-22268
15	154	9.4	1430	20	US-10-425-115-35659
16	153	9.4	1294	19	US-10-437-963-34215
17	148.5	9.1	981	17	US-10-369-493-37730
18	148.5	9.1	1545	20	US-10-425-115-131728
19	146.5	9.0	1361	19	US-10-437-963-40955
20	146.5	9.0	1425	18	US-10-424-599-115510
21	145	8.9	1453	19	US-10-437-963-78719
22	144.5	8.8	1447	19	US-10-437-963-46106
23	143.5	8.8	1344	19	US-10-767-701-13412
24	143	8.7	1140	15	US-10-136-444-7
25	140.5	8.6	1287	19	US-10-437-963-74437
26	139.5	8.6	1387	18	US-10-425-114-25033
27	139.5	8.5	1407	20	US-10-425-115-16262
28	139.5	8.5	1449	17	US-10-431-873-44
29	135.5	8.3	777	17	US-10-369-493-4166
30	135.5	8.3	1059	19	US-10-437-963-93028
31	134.5	8.2	1322	19	US-10-437-963-38404
32	133.5	8.2	1293	22	US-10-819-773-4
33	133.5	8.2	1345	16	US-10-171-174A-13
34	133.5	8.2	1392	18	US-10-424-599-48151
35	132	8.1	806	17	US-10-369-493-42893
36	131.5	8.0	927	17	US-10-369-493-31336
37	131	8.0	1613	18	US-10-425-114-27030
38	131	8.0	2666	20	US-10-739-930-1684
39	130.5	8.0	1165	18	US-10-425-114-35376
40	130.5	8.0	1332	18	US-10-425-114-27849
41	130.5	8.0	1332	20	US-10-425-115-82503
42	130.5	8.0	1375	18	US-10-424-599-68643
43	130	7.9	3914	19	US-10-437-963-9728
44	129	7.9	960	17	US-10-369-493-37707
45	128.5	7.9	1075	19	US-10-437-963-40100

ALIGNMENTS

RESULT 1
US-10-105-319-1
; Sequence 1, Application US/10105319
; Publication NO. US20030190695A1
; GENERAL INFORMATION:
; APPLICANT: Yun-Bor Yang
; APPLICANT: Chia-Li Wei
; APPLICANT: Jyh-Shing Hsu
; APPLICANT: Ying-Chieh Tsai
; TITLE OF INVENTION: Mutated Penicillin Expandase and Process for Preparing 7-ADCA Usi
; FILE REFERENCE: 6653-017-999
; CURRENT APPLICATION NUMBER: US/10/105,319
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 1

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(1164)
US-10-105-319-1

Alignment Scores:
Pred. No.: 1,89e-209 Length: 1230
Score: 1636.00 Matches: 311
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

10719236-1_232-1164 (1-311) x US-10-105-319-1 (1-1230)

Qy 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
Db 232 ATGACACAGACGGTGGCCACCTTCAGCTGGCCGAACTCCAGAGGGCTTCACACAGGAC 291
Qy 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
Db 292 GAGTTCGGCAGGTGTCTAGAGGACAGAGGCTCTTCTATCTGACGGACTGGGCTGTACC 351
Qy 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGluHisGlySerGlu 60
Db 352 GACACCGAGCTGAAGTCCGCGCAGAGACATCGTCACTTCGCTTCGAGCAGCGACGCGAG 411
Qy 61 AlaGluLysArgAlaValThrSerProValProThrMetLysArgGlyPheThrGlyLeu 80
Db 412 GCGAGAGACCGCGCGCTCACTCGCCGCTCCCAACATCGCGCGGCTTCACGAGGCTG 471
Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db 472 GAGTCGGAGACACCGCCCAATGATCCAAATACCGGCACTCTCCGACTACCTGATGTGC 531
Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
Db 532 TACTCGAAGGGCACCAGCGCGACCACTCTTCCGTCGGTGACTTCGAGCGGATCTGGACC 591
Qy 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db 592 CAGTACTTCGACCGCCAGTACACGCTCCGCGGGGTGCGCCGGAGGATCTCTGGGGCG 651
Qy 141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db 652 ACCGGGACCGAGCCCGAGGGGGGTGAGGCTTCTTCGACTGCGAGCCGCTGTGCGG 711
Qy 161 PheArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMet 180
Db 712 TTCGGTACTCTCCGCAAGTCCCGAGACACCGAGCGCGGAGAGCCCTCGGATG 771
Qy 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db 772 GCGCGGACCTAGACCTGTCAATGATCACCTTCATCCAGACACACCTTGCGCCAAAGGC 831
Qy 201 PheValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeuProTyrArgProAsp 220
Db 832 TTCGTCAAGCTCCAGGCGCGAGGTGGCGGGCGTTCAAGACCTGCTTACGTCGGAC 891
Qy 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValValLysAla 240
Db 892 GCCGCTCCCTCTCTCTGCGGGCGCATTCGACCCCTGGTGAACGGCGGCGCAAGGCC 951
Qy 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db 952 CCGCGGACCATGTGCGGGCCCCCGCAGGAGACAGATAGCGGGAGAGCGGACCTTCC 1011
Qy 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
```

```

|||||
Db 1012 AGTGTGTTCTTCTCCGTCCTCCCAAGCGGACTTACACTTCCGTCGGCGGCGGAG 1071
Qy 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGlyGly 300
Db 1072 TGCAGCTTCATGTCAAGCTCGAGCGGCGAGACCGCCACGTTCCAGATGATCGGGGGC 1131
Qy 301 AsnTyrValAsnIleArgArgThrSerLysAla 311
Db 1132 AACTAGTGAACATCGCGCCGACATCCAGGCA 1164

RESULT 2
US-10-719-237-1
; Sequence 1, Application US/10719237
; Publication No. US20040077034A1
; GENERAL INFORMATION:
; APPLICANT: Yunn-Bor Yang
; APPLICANT: Chia-Li Wei
; APPLICANT: Jyh-Shing Hsu
; APPLICANT: Ying-Chieh Tsai
; TITLE OF INVENTION: Mutated Penicillin Expandase and Process for Preparing 7-ADCA Unit
; FILE REFERENCE: 6653-017-999
; CURRENT APPLICATION NUMBER: US/10/719,237
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/10/105,319
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(1164)
US-10-719-237-1

Alignment Scores:
Pred. No.: 1,89e-209 Length: 1230
Score: 1636.00 Matches: 311
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

10719236-1_232-1164 (1-311) x US-10-719-237-1 (1-1230)

Qy 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
Db 232 ATGACACAGACGGTGGCCACCTTCAGCTGGCCGAACTCCAGAGGGCTTCACACAGGAC 291
Qy 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
Db 292 GAGTTCGGCAGGTGTCTAGAGGACAGAGGCTCTTCTATCTGACGGACTGGGCTGTACC 351
Qy 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGluHisGlySerGlu 60
Db 352 GACACCGAGCTGAAGTCCGCGCAGAGACATCGTCACTTCGCTTCGAGCAGCGGAGGAG 411
Qy 61 AlaGluLysArgAlaValThrSerProValProThrMetLysArgGlyPheThrGlyLeu 80
Db 412 GCGGAGAGAGCGCCGTCACCTCGCCGCTCCCAACATCGCGCGGCTTCACGAGGCTG 471
Qy 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
Db 472 GAGTCGGAGAGACACCGCCCAATGATCCAAATACCGGCACTCTCCGACTACCTGATGTGC 531
Qy 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
Db 532 TACTCGAAGGGCACCAGCGCGACCACTTCTCCGTCGGTGACTTCGAGCGGATCTGGACC 591
Qy 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
```



```
Db      592 CAGTACTTGACGCGCAGTACACCGGCTCCGGGGGTGGCCGGGAGGTCTCGGGGGG 651
Qy      141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db      652 ACCGGGACCGAGCCCGACCGCGGGGTGAGGCGCTTCTCGACTGCGAGCCGCTGTCGGG 711
Qy      161 PheArgTyrPheProGluValProGluHisArgSerAlaGluGluGluProLeuArgMet 180
Db      712 TTCGGCTACTTCCCGCAGGTCCTCCGAGCACCGGAGCCGAGAGAGCCCTCGCGGATG 771
Qy      181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db      772 GCGCGGCATTAAGACCTGTGCATGGTCACTTCATCCACAGACACCTCGCGCCACAGGCG 831
Qy      201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db      832 TTCGTCAGCTTCAGGCGCCAGGTCGGCGGCTTCACGAGCCTGCTACCGTCCGGGAC 891
Qy      221 AlaValIleuValPheCysGlyValAlaIleThrLeuValThrGlyGlyValIleValAla 240
Db      892 GCGGCTCTGCTCTTCTCGCGGCGCATCGGACCTGGTGACCGGCGGAGGTCAGAGGCG 951
Qy      241 ProArgHisHisValAlaAlaIleProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db      952 CCCCGGCACATGTCGGCGGCGCCCGCGAGAGACAGATAGCGGGGAGCGAGCGCACTCC 1011
Qy      261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db      1012 AGTGTGTTCTTCTCTCGCTCCCAACGGGAGCTTCACCTTCTCGCTCCGCTGGCGGGGAG 1071
Qy      281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGluAspTrpIleGlyGly 300
Db      1072 TCCGGCTTCGATGTCAGGCTCGGAGCGGCGAGCCGACCTTCAGGATTGATCGGGGCG 1131
Qy      301 AsnTyrValAsnIleArgArgThrSerIleValAla 311
Db      1132 AACTACGTGAACATCCGCGGACATCCAGGCA 1164

RESULT 3
US-10-719-236-1
; Sequence 1, Application US/10719236
; Publication No. US2004007035A1
; GENERAL INFORMATION:
; APPLICANT: Yunn-Bor Yang
; APPLICANT: Chia-Li Wei
; APPLICANT: Jyh-Shing Hsu
; APPLICANT: Ying-Chieh Tsai
; TITLE OF INVENTION: Mutated Penicillin Expandase and Process for Preparing 7-ADCA Usi
; FILE REFERENCE: 6653-017-999
; CURRENT APPLICATION NUMBER: US/10/719,236
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/10/105,319
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)..(1164)
US-10-719-236-1

Alignment Scores:
Pred. No.: 1,896-209          Length: 1230
Score: 1636.00               Matches: 311
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00%         Indels: 0
DB: 18                        Gaps: 0
```

```
10719236-1_232-1164 (1-311) x US-10-719-236-1 (1-1230)
Qy      1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAsp 20
Db      232 ATGGACAGACGAGTGCCCACTTCCTGAGCTGGCCCAACTCCAGCAGGGGCTCCACACAGAG 291
Qy      21 GluPheAspArgCysIleuArgAspIleGlyLeuPheTyrIleuThrAspCysGlyIleuThr 40
Db      292 GAGTTCGCGAGGTGTCTGAGGAGCAAGGCTCTTCTATCTGACGAGACTGTGGTCTGACC 351
Qy      41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPhePheGluHisGlySerGlu 60
Db      352 GACACCGAGCTGAGAGTCGGCCCAAGGACATCGTCAATCGACTTCTTCGAGCAGGAGGAGAG 411
Qy      61 AlaGluYsArgAlaValThrSerProValProThrMetArgArgIlePheThrGlyLeu 80
Db      412 GCGGGAAGCCCGCGCTGACCTCCGCTCCCAATGCGCGCGCTTCACCGGCTTCACCGGGCTG 471
Qy      81 GluSerGluSerThrAlaGlnIleThrAspThrGlySerTyrSerAspTyrSerMetCys 100
Db      472 GAGTGGAGACACCGCCCAATCAATACCGGAGCTTACTCGACTTACTCGATGTGC 531
Qy      101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTrpThr 120
Db      532 TACTCGATGGGCACCGGGGACCAACTCTTCCGTCGGTGACTTCGAGCGGATCTGGAGCC 591
Qy      121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValIleuArgAla 140
Db      592 CAGTACTTGACGCGCCAGTACACCGCTCCCGCGGCTCCGCGGAGTCTTCGGGGGCG 651
Qy      141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuLeuArg 160
Db      652 ACCGGGACCGAGCCCGAGCGGGGTGAGGCTTCTCGACTGGAGCCCTGTGTGGG 711
Qy      161 PheArgTyrPheProGluValProGluHisArgSerAlaGluGluGluProLeuArgMet 180
Db      712 TTCGGCTACTTCCCGCAGGTCCTCCGAGAGCCGAGCGCGGAGAGCCCTCGGGATG 771
Qy      181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db      772 GCGCGGCATTAAGACCTGTGCATGGTCACTTCATCCAGACACACCTCGGCCACAGGCG 831
Qy      201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyrArgProAsp 220
Db      832 TTCGTCAGCTTCAGGCGCCAGGTCGGCGGCTTCACGAGCCTGCTACCGTCCGAGC 891
Qy      221 AlaValIleuValPheCysGlyValAlaIleThrLeuValThrGlyGlyValIleValAla 240
Db      892 GCGGCTCTGCTCTTCTCGCGGCGCATCGGACCTGGTGACCGGCGGAGGTCAGAGGCG 951
Qy      241 ProArgHisHisValAlaAlaIleProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db      952 CCCCGGCACATGTCGGCGGCGCCCGCGAGAGACAGATAGCGGGGAGCGAGCACTCC 1011
Qy      261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db      1012 AGTGTGTTCTTCTCGCTCCCAACGGGAGCTTCACCTTCTCGCTCCGCTGGCGGGGAG 1071
Qy      281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGluAspTrpIleGlyGly 300
Db      1072 TCCGGCTTCGATGTCAGGCTCGGAGCGGCGAGCCGCAAGCTTCAGGATTGATCGGGGCG 1131
Qy      301 AsnTyrValAsnIleArgArgThrSerIleValAla 311
Db      1132 AACTACGTGAACATCCGCGGACATCCAGGCA 1164

RESULT 4
US-10-719-238-1
; Sequence 1, Application US/10719238
; Publication No. US20040121425A1
; GENERAL INFORMATION:
; APPLICANT: Yunn-Bor Yang
```

```

; APPLICANT: Chia-Ii Wei
; APPLICANT: Jyh-Shing Hsu
; APPLICANT: Ying-Chieh Tsai
; TITLE OF INVENTION: Mutated Penicillin Expandase and Process for Preparing 7-ADCA Usi
; FILE REFERENCE: 6653-017-999
; CURRENT APPLICATION NUMBER: US/10/719,238
; PRIORITY FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/10/105,319
; PRIORITY FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(1164)
US-10-719-238-1

Alignment Scores:
Pred. No.: 1,896-209 Length: 1230
Score: 1636.00 Matches: 311
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

10719236-1_232-1164 (1-311) x US-10-719-238-1 (1-1230)
QY 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAsp 20
DB 232 ATGACACGACGCGTCCCACTTCAGCTGGCCGAATCCAGCAGGCGCTGCACACGAGAC 291
QY 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
DB 232 GAGTCCCGCAGGTGCTGTAGGAGCAAGGCGCTCTTCTATCTGACGAGATGGCGGTGACC 351
QY 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySerGlu 60
DB 352 GACACCGAGCTGAATCGCGCCCAAGACATCTCATCGACTTCTTCAGACACGCGCAGG 411
QY 61 AlaGluLysArgAlaValThrSerProValProThrMetCysArgGlyPheThrGlyLeu 80
DB 412 GCGGAGAACGGCGCTGCACTCGCCGCTCCCACTACATGCGCGGCTTCAACGAGCTG 471
QY 81 GluSerGluSerThrAlaGlnIleThrAsnThrGlySerTyrSerAspTyrSerMetCys 100
DB 472 GAGTCGGAGAGACCGCCCAATCAATACCGGCACTACTCGACTCATCGATGTC 531
QY 101 TyrSerMetGlyThrAlaAspAsnLeuPheProSerGlyAspPheGluArgIleTyrThr 120
DB 532 TACTCGATGGGACCGCGGACCACTCTTCCGCTCGGTGACTTCGACGCGATCTGGACC 591
QY 121 GlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArgAla 140
DB 592 CAGTACTTCGACCGCGCAGTACCGCGCTCCGCGGGGTGCGCGGAGAGTCTTCGCGGCG 651
QY 141 ThrGlyThrGluProAspGlyGlyValAlaGluAlaPheLeuAspCysGluProLeuLeuArg 160
DB 652 ACCGGGACCGAGCCCGACGCGCGGGTTCGAGGCTTCTCTGACTGAGACCGCTGTGGGG 711
QY 161 PheArgTyrPheProGlnValProGluHisArgSerAlaGluGluGlnProLeuArgMet 180
DB 712 TTCCTCTCTTCCCGAGGTCCCGACGACCGCGGCGGAGACACCTCGCGGATG 771
QY 181 AlaProHisTyrAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
DB 772 GGGCGGACTAGACCTGTGATGGTCACTTCATCCAGACACACCTCGCGCCACAGGC 831
QY 201 PheValSerLeuGlnAlaGluValGlyAlaPheThrAspLeuProTyrArgProAsp 220
DB 831
```

```

DB 832 TTGCTACGCTCCAGGCGAGGTCGGGCGCGGCTTACAGGACCTGCCTACGTCGGAC 891
QY 221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValIleValAla 240
DB 892 GCGTCTCTCTTCTGCGCGGCGCATCCGACCTGTGACCGCGCGCCAGGTCAAGGCC 951
QY 241 ProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
DB 952 CCGGCGACCATGTGCGGCGCCCGCGGAGCACCAATACGGGCGAGCGCACCTTC 1011
QY 261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
DB 1012 AGTGTTCTTCTCTCGCTCCGACCAAGCGGACTTCCTTCGCTCCGCTGCGCGCGAG 1071
QY 281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyrIleGlyGly 300
DB 1072 TCGGCTTCATGTACCTGACCTGAGCGGAGACCGCACCGTTCAGAGATTGATGGGGGC 1131
QY 301 AsnTyrValAsnIleArgArgThrSerLysAla 311
DB 1132 AACTACTGAAATCCGCGCGCAATCCAGGCA 1164

RESULT 5
US-10-789-071-1
; Sequence 1, Application US/10789071
; Publication No. US20040259191A1
; GENERAL INFORMATION:
; APPLICANT: Yunn-Bor Yang
; APPLICANT: Chia-Ii Wei
; APPLICANT: Jyh-Shing Hsu
; APPLICANT: Ying-Chieh Tsai
; TITLE OF INVENTION: Mutated Penicillin Expandase and Process for Preparing 7-ADCA Usi
; FILE REFERENCE: 6653-017-999
; CURRENT APPLICATION NUMBER: US/10/789,071
; PRIORITY FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US/10/105,319
; PRIORITY FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(1164)
US-10-789-071-1

Alignment Scores:
Pred. No.: 1,896-209 Length: 1230
Score: 1636.00 Matches: 311
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

10719236-1_232-1164 (1-311) x US-10-789-071-1 (1-1230)
QY 1 MetAspThrThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAsp 20
DB 232 ATGACACGACGCGTCCCACTTCAGCTGGCCGAATCCAGCAGGCGCTGCACACGAGAC 291
QY 21 GluPheArgArgCysLeuArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeuThr 40
DB 292 GAGTCCCGCAGGTGCTGTAGGAGCAAGGCGCTCTTCTATCTGACGAGATGGCGGTGACC 351
QY 41 AspThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySerGlu 60
DB 352 GACACCGAGCTGAATCGCGCCCAAGACATCTCATCGACTTCTTCAGACACGCGCAGGAG 411
QY 61 AlaGluLysArgAlaValThrSerProValProThrMetCysArgGlyPheThrGlyLeu 80
DB 411
```

```
Db      412 GCGAGAACGCGCGCTGCACCTGCCCTGCCACCATTGCGCGGCTTACCGGCGCTG 471
Qy      81 GIUSenGISeThrAlaGlnIleThraThrngISeTySerAspTySerMetCys 100
Db      472 GAGTGGAGAGACCGCCGAGATCACCAATACCGGAGCTACTCCGACTACCTGATGTC 531
Qy      101 TySerMetGIYThraAlaAspAsnLeuPheProSerGIYAspPheGluArgIleTrpThr 120
Db      532 TACTGATGGGACCGCGGACCAACCTCTCCGTCGGGAGTCTTCAGCGGAGATCTGACC 591
Qy      121 GlnTyPheAspArgGlnTyThraAlaSerArgAlaValAlaArgGluValLeuArgAla 140
Db      592 CAGTCTTCGACCGGACAGTACACCGCTCCCGCGCGTCCGCGGAGAGTCTGCGGCG 651
Qy      141 ThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeuArg 160
Db      652 ACCGGAGACCGAGCCGAGCGGCGGTCGAGGCTTCTCTGACTGCGAGCGCTGTCGG 711
Qy      161 PheArgTyPheProGlnValProGluHisArgSerAlaGluGluProLeuArgMet 180
Db      712 TTCCTGCTACTTCCCGAGGTCCCGAGCACCCGAGCGCGGAGAGCAGCCCTGCGGATG 771
Qy      181 AlaProHisTyPheAspLeuSerMetValThrLeuIleGlnGlnThrProCysAlaAsnGly 200
Db      772 GCGCGGACACTACGACTGTGATGTGACCTCATCTCAGACAGACACCTGCGCCAAACGCG 831
Qy      201 PheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAspLeuProTyArgProAsp 220
Db      832 TTCGTAGCTTCAGAGCCGAGGTCCGCGCGGCTTTCACGAGCTCCCTACCGCTCGAGC 891
Qy      221 AlaValLeuValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValValLeuAla 240
Db      892 GCGGCTCTGCTTCTTCTGGGCGCCCATCGGACCTGTGATCAGCGGCGGAGTCAAGGCC 951
Qy      241 ProArgHisAlaValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSer 260
Db      952 CCCCCGACATGTCGCGGCGCCCGCGAGGACAGATGAGCGGAGCAGCGCACCTCC 1011
Qy      261 SerValPhePheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGlu 280
Db      1012 AGTGTGCTTCTCTCCGTCACGCGGACTTCACCTTCTCGTCCGTCGCGGCGGAG 1071
Qy      281 CysGlyPheAspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrpIleGlyGly 300
Db      1072 TCCGCTTCGATGTGAGCTGAGCGGAGACCGGACCGCAGCTTCAAGATTGATCGGAGC 1131
Qy      301 AsnTyValAsnIleArgArgThrSerIysAla 311
Db      1132 AACTACGTAAACATCCGCGCACATCCAAGGCA 1164

RESULT 6
US-10-369-493-35926
; Sequence 35926, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIORITY FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIORITY FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO: 35926
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35926
```

```
Alignment Scores:
Pred. No.: 7 31e-12 Length: 1101
Score: 169.00 Matches: 85
Percent Similarity: 36.63% Conservation: 41
Best Local Similarity: 24.71% Mismatches: 158
Query Match: 10.33% Indels: 60
DB: 17 Gaps: 14

10719236-1_232-1164 (1-311) x US-10-369-493-35926 (1-1101)
Qy      1 MetAspThrThraValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGln--- 19
Db      82 ATGCCCAATAATAGTGGCTGTGCTGATGATGAGCCGCTGTGAACAGGAGCGGCTCGAAGAC 141
Qy      20 -----AspGluPheArgArgCysLeuArgAspGlyLeuPheTyThr 35
Db      142 CGGACATTTCTAGGCGACTTCGTCCGCTCGCGGACATCGGCTTTTCTATCTTCC 201
Qy      36 AspCysGlyLeuThrAspThrGluLeuIlySerAlaIlyAspIleValIleAspPhe 55
Db      202 GACATGGCATCTCTGGCGAGATACGAGAGTGTCTACCGCATCGCGCAGTCTTT 261
Qy      56 GluHisGlySerGluAlaGluIlyAspAlaValThrSerProValProThrMetArgArg 75
Db      262 GCTTGG---CCGAGGCGGACAAAGCTGGCATTCGAGATGTCAGATCTCTCCAGTTTCC 318
Qy      76 GlyPheThrGlyLeuGluSerGluSerThraAlaGlnIleThraThrGlySerTySer 95
Db      319 GGTATACGCGCGCGCGGCGGAGAACTGACTAG-----GGCAGAGAG 360
Qy      96 AspTySerMetCysTyPheSerMetGlyThrAlaAspLeuPheProSerGlyAspPhe 115
Db      361 GACTGGCGGCAACGCTTGACATGCGGTGAGCGGCGGAGCAATCGGCAAGCGCGGGA 420
Qy      116 GluArgIleTrpThrGln---TyrPheAspArgGlnTyThraAlaSer----- 130
Db      421 ACTCCCGCTGACCCCGCTCCAGACCGCAACAGTGGCCGCGGCTCCCGGATTC 480
Qy      131 -----ArgAlaValAlaArgValLeuArg 139
Db      481 AAACCGCGCTCTTGGCTCGGACAGACAGGTGACACCGCGGCGGCTCTCTGAG 540
Qy      140 AlaThrGlyThrGluProAspGlyGlyValGluAlaPheLeuAspCysGluProLeu 159
Db      541 GCTTCCGCGCATGCTCTGATCAGCCCGAGGATCGTTC-----GATCGATC--- 588
Qy      160 ArgPheArgTyPheProGlnValProGluHisArg---SerAlaGluGluProLeu 178
Db      589 -----TACACGAGGAGCCCAACCATGCGATGAAGATCGCTACCCCGC 636
Qy      179 Arg-----MetAlaProHisTyPheAspLeuSerMetValThr 190
Db      637 CGTGACCAACCGCGCGGACGAGTGGCGGACGAGCAAGAGTGGTTCCTCAC 696
Qy      191 LeuIleGlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGly 210
Db      697 CTGCTGCTTCAG-----GATGACAAATAGGCGCTGCGATGATGACGCG 744
Qy      211 AlaPheThrAspLeuProTyArgProAspAlaValLeuValPheCysGlyAlaIleAla 230
Db      745 AGCTGGGATGTGAGACCCATTCCTCCGACGCTGTCTGTCACATCGCGGAATTGCTC 804
Qy      231 ThrLeuValThrGlyGlyValIlyAspAlaProArgHisIleValAlaAlaProArgArg 250
Db      805 GAACGTGCGCTCAACGCTACTTCCGCGGACCGTGCATGCTGTCGAGACCGCC--- 861
Qy      251 AspGlnIleAlaGlySerSerArgThrSerSerValPhePheLeuArgProAsnAlaAsp 270
Db      862 -----GCCGCGTGCAGCGCATTTCCGTGCGCTTCTTTCACGCGGCGGCTGAT 912
Qy      271 PheThr-----PheSerValProLeuAlaArgGluCysGlyPheAspVal 285
```

```

Db      913 GCGACGATACCGCTGCTTCCGAGGAACCTGGCGGCGACGGCGCGCGGCGCA 972
Qy      286 SerLeuaspGlygluThrAlaThrPhegluAspTrpIleGlyValAsnIle 305
Db      973 AGC---GATCCGGAACATCCGCTGTTCCGAT---GTGGAAACCAATGTCTCAAAAGC 1026
Qy      306 ArgArgThrSer 309
Db      1027 GCGCTGCGCTCG 1038

RESULT 7
US-10-494-495-1
; Sequence 1, Application US/10494495
; Publication No. US20050079498A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham plc
; APPLICANT: The Governors of the University of Alberta
; TITLE OF INVENTION: Polynucleotides and Polypeptides Involved in Clavulanic
; TITLE OF INVENTION: Acid Biosynthesis and Use Thereof
; FILE REFERENCE: P32913
; CURRENT APPLICATION NUMBER: US/10/494,495
; PRIOR FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: PCT/GB02/04989
; PRIOR FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: GB 0126756.6
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: GB 0128776.2
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 29870
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
US-10-494-495-1

Alignment Scores:
Pred. No.:      1.32e-09      Length:      29870
Score:          168.00      Matches:      77
Percent Similarity: 39.32%      Conservative: 50
Best Local Similarity: 23.84%      Mismatches: 150
Query Match:    10.27%      Indels:      46
                        Gaps:      14
10719236-1_232-1164 (1-311) x US-10-494-495-1 (1-29870)

Qy      5 ValProThrPheSerLeuAlaGluLeu-----GlnGlnGlyLeuHis 18
Db      1307 GTTCCGACCATCGACATCTCGCCGCTGTTCCGAAACGACGCCGCGAAGAACCGCGTC 1366
Qy      19 GlnAspGluPheArgArgCysLeuArgAspIleGlyLeuPheTyLeuThrAspCysGly 38
Db      1367 GCCAGAGGATACACAGGGGCTCGCCGCGCTTCTTCTACGCCACGACACCGAGC 1426
Qy      39 LeuThrAspTrpGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGly 58
Db      1427 GTG-----GACGTCCAGAGCTCCAGACCGGTGGGAACGAGTTCCACGGGCGCATG 1477
Qy      59 SerGluAlaGluLys-----ArgAlaValThrSerProValProThrMet 73
Db      1478 ACCGACCGAGAGAACGACGACCTCGCATCCACCGGTACCAACCGGACMACCCGACGTG 1537
Qy      74 ArgArgGlyPhe-----ThrGlyLeuGluSerGluSerThrAlaGlnIleThr 89
Db      1538 CGCAACGGCTTACTACAGCGCGTCCCGGAGAGAGCGCGCTTCGATCTTGTACTTC 1597
Qy      90 AsnThrGlySerTySerAspTySerMetCysTySerMetGlyThr-----Ala 106
Db      1598 AACCCGAGACTTCGGGAGGACACACCCGATG---ATCGCGGGGAGACGCCGATGACAGAG 1654
Qy      107 AspAsnLeuPheProSerGlyAsp-----PheGluArgIleTrpThrGlnIleTy 122
Db      1655 GTGAACCTTGCGCCGACGAGAGCGGACCCGCGCTTCCGCGCTTCTGCGAGGGCTAC 1714

```

```

Qy      123 PheAspArgIleGlnTrpThrAlaSerArgAlaValAlaArgGluValLeuArgIleTrpGly 142
Db      1715 TACCGGAGAGTGTGAAGACTCTCACCGTCTCATGCGGGGCTGGCGCTCGGCGG 1774
Qy      143 ThrGluProAspGlyGlyValGluAla-----PheLeuAspCysGlu 156
Db      1775 ---AGCGCGAGACATCTTCACGCGGCGGCTCGCGGAGCAGGACCTCGTGTCCGTC 1831
Qy      157 ProLeuLeuArgPheArgIleGlyTrpPheProGlnValProGluHisArgSerAlaGluGlu--- 175
Db      1832 TCCTGTATCCGCTACCCGCTATCTGAGAGAGTACCCCGGTGAAGACGGGTCCGACGCG 1891
Qy      176 GlnProLeuArgMetAlaProHisIleTyArgPheSerMetValThrLeuIleGlnIleTrp 195
Db      1892 CAGCTCTGAGCTTCGAGACGACCTTCGAGCTCTGATCTCATGTCACCGGTCTTCACAGACC 1951
Qy      196 ProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyValAlaPheThrAspLeu 215
Db      1952 CAGGTGCAGAAC-----CTCCAGGTGAGACGCGTCCGACGCGCTCGCGGACATC 1999
Qy      216 ProTyArgProAspAlaValLeuValPheCysGlyAlaIleAlaThrLeuValIleTrpGly 225
Db      2000 CCGACGTCGAGAGACGACTTCTGTCATCGGTACCTACATGCGCGCATGTCACAGAAC 2059
Qy      236 GlyGlnValLysAlaProArgHisHisValAlaAlaProArgArgAspGlnIleAlaGly 255
Db      2060 GACTACTCCCGCGCGCCCAACACCGGGTC-----AAGTTCCGGAAC 2101
Qy      256 SerSerArgThrSerSerValPhePheLeuArgProAsnAlaAspPheThrPheSerVal 275
Db      2102 GCGAGAGGCTGTCCCGCGCTTCTTCACGCGGCGACGAGCGGCGCATCGAG--- 2158
Qy      276 ProLeuAlaArgGluCysGlyPheAspValSerLeuAspGlyIleThrAlaThrPheGln 225
Db      2159 CCGTTCGTCCGAGAG---GGCGGACGAGAGGTAGAGAACGAGGCCCTGTCTACGGG 2215
Qy      296 AspTrpIle 298
Db      2216 GACTACTCTC 2224

RESULT 8
US-10-425-114-22510
; Sequence 22510, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-2A(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22510
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB587-225-H12_FLI
US-10-425-114-22510

Alignment Scores:
Pred. No.:      5.49e-10      Length:      1335
Score:          156.00      Matches:      78
Percent Similarity: 38.02%      Conservative: 49
Best Local Similarity: 23.35%      Mismatches: 119
Query Match:    9.54%      Indels:      88
                        Gaps:      16

```

10719236-1_232-1164 (1-311) x US-10-425-114-22510 (1-1335)

```
OY 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAspGluPheArg 23
DB 205 ACCTCAGGAGGAGGCTCTCCCTCCATCTGTGACCTCTCCCTGCGCGGACGAGGTGGCG 264
OY 24 ArgCysLeu-----ArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeu 39
DB 265 CCGCGCATCTCGAGGCGCGGACAGAGATCGCTTCTCCAGGTGTGAACACGCGCTC 324
OY 40 ThrAspThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySer 59
DB 325 TCCCTGAGAGCGCATGAGACATGAGACGCTGTGCGACAGATTCTTCGCG-----CTG 378
OY 60 GluAlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGly 79
DB 379 CCGCGCGAGACAGAGCC-----GGG 399
OY 80 LeuGluSerGluSerThrAlaGlnIleThr-----AsnThr 91
DB 400 CTGTACTCGAGAGACACGCGGAGGCGGACCGGATCTACTCCAGACCATGTTCCAGACG 459
OY 92 GlySerTyrSerAspTyrSerMetCysTyrTyrSerMetGlyThrAlaAspAsnLeuPhePro 111
DB 460 GCGCGCGAGAACTACTGGCGGCGACTGCTCCGCTGCGCTGCTGCTCC-----TTCCCG 510
OY 112 Ser---GlyAspPheGluArgIleTrp-----ThrGlnTyrPhe 123
DB 511 GCGCTCGGAGACAGCGCGCGCTGTGCGGACAAAGCCCGGAGGCTCGCGGAGGTGGTG 570
OY 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgValLeuArg-----139
DB 571 GAGAGGTTCACTGACGACACGCGCGGCTGTGGAGATGAGATCTCGCTGCTGTGCGAG 630
OY 140 AlaThrGlyThrGluProAsp-----GlyGlyValGluAlaPhe 152
DB 631 GGCCTCGGCTCGCGCGGACTACTGAGAGGAGCATGACGAGGCGG-----678
OY 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGluHisArgSer 172
DB 679 -----GACGTGCTCTCCACGCTCAACCACTACCGGCTGTGCGGACCCCAACGCG 729
OY 173 AlaGluGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192
DB 730 ACG-----CTCGGCTCGCGCGGCTGTGACCGGACCTCTCACTTGTCTC 777
OY 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyGlyAlaPhe 212
DB 778 -----CTCCCGACGATGTGCTGCTGCG-----798
OY 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223
DB 799 CTGCGAGTCTGCTTACAGGCGGCTGATCAGAGGTGAGCGCTGTGCGCGCTTCTGCTC 858
OY 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValIleAspAlaProArgHis 243
DB 859 GTCACTTCGATGCCAGCTTGAAGTTGTGACAAAGGATTTGAAGAGCATCGACGAC 918
OY 244 HisValAlaIleAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
DB 919 CCGGTGATGACCAAC-----CTGGAGATGGCGCGGACAAACAGTGTGCGACG 963
OY 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283
DB 964 TTCATCATGCCACACACGAGCTGCTCATCGCG-----CCCGCGCGAGTTCTCTGACGAC 1020
OY 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrp 297
DB 1021 GACAAACCGCGCTGCTACCGGACCTCTACCTTGGCGGACTTC 1062
```

RESULT 9
US-10-425-115-35660

```
/ Sequence 35660, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plasmids
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO: 35660
/ LENGTH: 1559
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(1559)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MFT4577_132519C.1
/ US-10-425-115-35660

Alignment Scores:
Pred. No.: 6.9e-10 Length: 1559
Score: 156.00 Matches: 78
Percent Similarity: 38.02% Conservative: 49
Best Local Similarity: 23.35% Mismatches: 119
Query Match: 9.54% Indels: 88
DB: 20 Gaps: 16

10719236-1_232-1164 (1-311) x US-10-425-115-35660 (1-1559)
OY 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlyLeuHisGlnAspGluPheArg 23
DB 288 ACCTCAGGAGGAGGCTCTCCCTCCATGTCGACCTCTCCCTGCGCGGACGAGGTGGCG 347
OY 24 ArgCysLeu-----ArgAspLysGlyLeuPheTyrLeuThrAspCysGlyLeu 39
DB 348 CCGCGCATCTCGAGGCGCGGACAGAGATCGCTTCTCCAGGTGTGAACACGCGCTC 407
OY 40 ThrAspThrGluLeuLysSerAlaLysAspIleValIleAspPheGluHisGlySer 59
DB 408 TCCCTGAGAGCGCATGAGACATGAGACGCTGTGCGACGAGGTCTTCCG-----CTG 461
OY 60 GluAlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGly 79
DB 462 CCGCGCGAGACAGAGCC-----GGG 482
OY 80 LeuGluSerGluSerThrAlaGlnIleThr-----AsnThr 91
DB 483 CTGTACTCGAGAGACAGCGGAGGCGGACCGGATCTACTCCAGACCATGTTCCAGACG 542
OY 92 GlySerTyrSerAspTyrSerMetCysTyrTyrSerMetGlyThrAlaAspAsnLeuPhePro 111
DB 543 GCGCGCGAGAACTACTGGCGGCTGCTCCGCTGCGCTGCTGCTCC-----TTCCCG 593
OY 112 Ser---GlyAspPheGluArgIleTrp-----ThrGlnTyrPhe 123
DB 594 GCGCGCGGACGACGCGCGGCTGCTGCGGACGAGCGGAGGCTCGCGGAGGTGGTG 653
OY 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgValLeuArg-----139
DB 654 GAGAGTTCACTGACGACGCGGCGGCTGTGAGATCTCGCGCTGCTGCTGCGAG 713
OY 140 AlaThrGlyThrGluProAsp-----GlyGlyValGluAlaPhe 152
DB 714 GGCCTCGGCTCGCGCGGACTACTGAGAGGAGACATGACGCGGCG-----761
OY 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGluHisArgSer 172
DB 761 -----
```

Db 762 -----GACGTGCTCTCCACGCTACACCACTACCCGCCCTGCCCCGACCCCAAGGCC 812

Qy 173 AlaGlulGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192

Db 813 ACG-----CTCGAGCTGCTCGCCGCGCATCTGTGACCCGCAACTCTTCACCTTCCTC 860

Qy 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnIleValGlyAlaPhe 212

Db 861 -----CTCCCGACGATGTCCTCGGC----- 881

Qy 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223

Db 882 CTCGAGGCGCCCTACAGGGCGCATGTGATCAGGGTGAGCTGTGCCCCGGGCTTCGTGC 941

Qy 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValValAlaProArgHis 243

Db 942 GTCAACTTCGGATGCCAGCTTGAGATTGTGACGACGGAGATCTGAAGAGATCGACGAC 1001

Qy 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263

Db 1002 CCGGTGATGACCAAC-----CTGGAGTGGCGCGGACACACGCTGGCCACG 1046

Qy 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283

Db 1047 TTCATCATGCCACACGACGACTGCTCATCGAC---CCCCGCGCGAGTTCTCGAGCGAC 1103

Qy 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrp 297

Db 1104 GACAACCCGCGCTGCTACCCGACCTTCACCTTCGCCGACCTTC 1145

RESULT 10

US-10-425-114-16150

; Sequence 16150, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425.114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 16150

; LENGTH: 1256

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3062-030-F8_FLI

US-10-425-114-16150

Alignment Scores:

Pred. No.: 6.84e-10 Length: 1256

Score: 155.00 Matches: 78

Percent Similarity: 38.02% Conservative: 49

Best Local Similarity: 23.35% Mismatches: 119

Query Match: 9.47% Indels: 88

DB: 18 Gaps: 16

10719236-1_232-1164 (1-311) x US-10-425-114-16150 (1-1256)

Qy 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGluPheArg 23

Db 123 ACCTCAGGAGCGGTCTCCCTGCCCATCGTGACCTCTCCCTCGGCGCGGAGAGGTGGG 182

Qy 24 ArgCysLeu-----ArgAspGlyGlyLeuPheTyrLeuThrAspCysGlyLeu 39

Db 183 CGCGCATCTCGAGCGCGGACGAGGATCGCTTCTTCAGGTGTGAACACGAGCGCTC 242

Qy 40 ThrAspThrGluLeuLysSerAlaLysAspIleValIleAspPheGlnHisGlySer 59

Db 243 TCCCTGAGGCGGTGAGAGCATGAGACGGTGTGCCAGAGTTCTTCGG-----CTG 296

Qy 60 GluAlaGluLysArgAlaValThrSerProValProThrMetArgArgGlyPheThrGly 79

Db 297 CCGCGGAGGACAAAGCC-----GGG 317

Qy 80 LeuGluSerGluSerThrAlaGlnIleThr-----AsnThr 91

Db 318 CTCTACTCGAGGACACGGGAGGGGACCCGAGTCTACTCCAGCACCATGTTGCACAG 377

Qy 92 GlySerTyrSerAspTyrSerMetCysTyrSerMetGlyThrAlaAspLeuPhePro 111

Db 378 GAGCGCGAGAAAGTCTGGCGGCACTGCTTCGCTGCTGCTCC-----TTCCCC 428

Qy 112 Ser-----GlyAspPheGlnArgIleTyr-----ThrGlnTyrPhe 123

Db 429 GCCGTGCGGACAGCCCGCGGCTGCGCCGACACCCCGAGAGCTCCGGAGGTGTG 488

Qy 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArg----- 139

Db 489 GAGAGGTTCACGGTGCAGACGCGGGGCTGGGAGTGAATCTGTGCGCTGCTGTGCGAG 548

Qy 140 AlaThrGlyThrGluProAsp-----GlyGlyValAlaGluAlaPhe 152

Db 549 GGCCTGGGCTCGGCCCTCCGACTCTGGAAGGGAGCATCAGCGCGGC----- 596

Qy 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGlnHisArgSer 172

Db 597 -----GACGTGCTCTCCAGCTACCAACCACTACCTCCGCTGCCGACCCCAAGCC 647

Qy 173 AlaGlulGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192

Db 648 ACG-----CTCGAGCTGCTCGCCGCGCATCTGTGACCCGCAACTCTTCACCTTCGCTC 695

Qy 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyAlaPhe 212

Db 696 -----CTCCCGACGATGTCCTCGGC----- 716

Qy 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223

Db 717 CTCGAGGTCCCTACAGGGCGCATGTGATCAGGGTGAGCTGTGCCCCGGGCTTCGCTC 776

Qy 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValValAlaProArgHis 243

Db 777 GTCAACTTCGGATGCCAGCTTGAGTTGTGACGACGGATCTTGAAAGACATCGAGCAC 836

Qy 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263

Db 837 CCGGTGATGACCAAC-----CTGGAGTGGCGCGGACGACGATCGGCCACG 881

Qy 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283

Db 882 TTCATCATGCCACACGACGACTGCTCATCGGC---CCCCGCGCGAGTTCTTCAGCGAC 938

Qy 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrp 297

Db 939 GACAACCCGCGGTGCTACCGCACCTTCACCTTCGGGACCTTC 980

RESULT 11

US-10-437-963-43792

; Sequence 43792, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21 (53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 43792
 LENGTH: 1281
 TYPE: DNA
 ORGANISM: *Oryza sativa*
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_46914C.1
 US-10-437-963-43792

Alignment Scores:
 Pred. No.: 8.22e-10 Length: 1281
 Score: 154.50 Matches: 70
 Percent Similarity: 38.49% Conservative: 52
 Best Local Similarity: 22.08% Mismatches: 158
 Query Match: 9.44% Indels: 37
 DB: 19 Gaps: 10

10719236-1_232-1164 (1-311) x US-10-437-963-43792 (1-1281)

QY	5	ValProthPhSerleuAlaGluLeuGlnGlyLeuHiSerIn-----AspGluPhe	22
DB	246	ATTCGGTGATGACCTCGCGAGCTGCAGCGCGGCGGCGAGGACGAGCTCGACAACTTC	305
QY	23	ArgAcGcYsLeuAlaArgAspIsgIyLeuPheTyLeuThrAspCySgIyLeuThAAspThr	42
DB	306	AGGCTCGCCTCGGAGCAGATGGGGCTCTTCACAGTGTGAACCTGGAGATGGAGAGAG	365
QY	43	GluLeuYsSerAlaYsAspIleValIleAspPheGluHisIsgIySerGluAlaGlu	62
DB	366	ACGATGAGAGAGATGGAGAAAGCGGCGGAGGAGATTCTTATGCTGCCGCTGGAGAGAG	425
QY	63	LysAlaGlaValThrSerProValProThrMetArgGlyPheThrGlyLeuGluSer	82
DB	426	GAGAGATGACCCCATGAGGCC-----GGCGGCATTCACGGC	461
QY	83	GluSerThrAlaGlnIleThrAsnThrArgIserTyrSerAspTyrSerMetCysTyrSer	102
DB	462	TACGGCCACGCTTCCTCTTCTCCGACGACGACGAGCTGCACTGGTGCACAACTCTGCC	521
QY	103	MetGlyThr-----AlaAspAsnLeuPheProSer-----GlyAsp	114
DB	522	CTGGCGCTCGAGCCCGGCTTCATCCGCGCGCCCACTCTGCGCGACAACTCCGCGCAC	581
QY	115	PheGluArgIleTrrThrGlnTyrPheAspArgGlnTyrThrAlaSerArgAlaValAla	134
DB	582	TTTATTAAGACCGCTGAGGAATGACTCGAGTGAATCAGGAGCTGTCTCGCGCTCTG	641
QY	135	ArgGluValLeuAlaArgAlaThrGlyThrGluProAspGlyValGluAlaPheLeuAsp	154
DB	642	GAGCAATCGCGCGCGCTGGGGCTGGCGCGCCGCGAGGCTGAACGGGATGTTCTGGGAG	701
QY	155	CysGluProLeuAlaArgPheArgTyrPheProGlnValProGluHisArgSerAlaGlu	174
DB	702	GCGGTCGACGCGGTGAGATGAACCTTCTACCCCGCGTCCCGCGG-----CCG	749
QY	175	GluGlnProLeuAlaArgMetAlaProHisTyrAspLeuSerMetValThrLeuIleGln	194
DB	750	GAGCTGAGCTGGGGCTGAGCGGACCTCGACGCGCGCGGTACCGGTCTCAGCAG	809
QY	195	ThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyGlyAlaPheThrAsp	214
DB	810	GACGCGCGCTTCGCGGGCTGCAGAGTCTCCGGGCG-----GGCGCGGCTGGTGGCC	863
QY	215	LeuProTyrArgProAspAlaValLeuValPheCysGlyValAlaIleAlaThrLeuValThr	234
DB	864	GTCATATCCCGTCCCGCGCGCTCTGCTCAAGTGGCGACACGCTCGAGGTCTCAACC	923
QY	235	GlyGlyGlnValLysAlaProArgHisHis--ValAlaAlaProArgAspGlnIle	253

Db	Sequence	Accession	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB
Db	924	AATGCAAGGTACAAAGAGCGTGGAGCAACCGGCGGTGGCCAGCGCGAGCAACGAC-----	977					
Qy	254	AlaGlySerSerArgThrSerSerValPhePheLeuArgProAsnAlaAspPheThrPhe	273					
Db	978	-----CGATGTCGCGTCTGACACTTCTAGCGCGCGCGCGATACAGACGTGCAGCTC	1025					
Qy	274	SerValProLeuAlaArgGluCysGlyPheAspValSerLeuAspGlyGluThrAlaThr	293					
Db	1026	GCG-----CCGCTTGGCCG-----GAGCTCTGTCGCCAGCGGGAGCGCGCGCAGG	1067					
Qy	224	PheGlnAspTrpIleGlyGlyAsnTrpValAsnIleArgArgThrSerLys	310					
Db	1068	TACCGAGCGTACACACGCGGCGAGTACAGCGCGGCACTACGTACCAAGCCGG	1118					
RESULT 12								
US-10-425-114-1868								
/ Sequence 1868, Application US/10425114								
/ Publication No. US2004003488B1								
/ GENERAL INFORMATION:								
/ APPLICANT: Liu, Jingdong								
/ APPLICANT: Zhou, Yihua								
/ APPLICANT: Kovalic, David K.								
/ APPLICANT: Screen, Steven E.								
/ APPLICANT: Tabaska, Jack E.								
/ APPLICANT: Cao, Yongwei								
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With								
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement								
/ FILE REFERENCE: 38-21(53313)B								
/ CURRENT APPLICATION NUMBER: US/10/425,114								
/ CURRENT FILING DATE: 2003-04-28								
/ NUMBER OF SEQ ID NOS: 73128								
/ SEQ ID NO 1868								
/ LENGTH: 1346								
/ TYPE: DNA								
/ ORGANISM: Zea mays								
/ FEATURE:								
/ OTHER INFORMATION: clone ID: 700169504_FLI								
US-10-425-114-1868								
Alignment Scores:								
Pred. No.: 1,036-09 Length: 1346								
Score: 154.00 Matches: 78								
Percent Similarity: 37.72% Conservative: 48								
Best Local Similarity: 23.35% Mismatches: 120								
Query Match: 9.41% Indels: 88								
DB: 18 Gaps: 16								
1071923								

Qy 112 Ser---GLYAspPheGluArgIleTrp-----ThrGlnTyrPhe 123
 Db 534 GCCCTGCGGACACGCGCGCGCTGGCGGACCAAGCCCGGAGCTCCGGAGGTGGTG 593
 Qy 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgValLeuArg----- 139
 Db 594 GAGAGGTTACAGGTGACAGCGCGGGGCTGGGATGAGATCTCGCTGCTGGAG 653
 Qy 140 AlaThrGlyThrGluProAsp-----GlyGlyValGluAlaPhe 152
 Db 654 GCGCTCGGCTCCGCGCGGACCTGGAGAGGACATCAGCGGGCGG----- 701
 Qy 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGlnHisArgSer 172
 Db 702 -----GACGTGCTCTCCAGCTCAACCACTACCGCGCTGCGCGCGCGCGCGCG 752
 Qy 173 AlaGluGlnGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192
 Db 753 ACG-----CTGGGCTGCGCGCGGACCTGTACCGCAACCTCTCACTTGGCTC 800
 Qy 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyGlyAlaPhe 212
 Db 801 -----CTCCCGACAGTGTACTGGC----- 821
 Qy 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223
 Db 822 CTCGAGGTCGCTTACGGGGGCGACTGATCCGGGTGAGCTGTGCGCGCGCTTGGCTC 881
 Qy 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValValAlaProArgHis 243
 Db 882 GTCAACTTCGGATGCGAGCTTGAGTGTGACGAACGGAGCTTGAAGAGCATGAGAC 941
 Qy 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
 Db 942 CGGGTGATGACCAAC-----CTGGAGTGGCGCGACACACGGTGGCCACG 986
 Qy 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283
 Db 987 TTCATCATGCCACACGAGCTGCTCATCGGC---CCGCGCGCGAGTTCTTCAAGGAC 1043
 Qy 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrp 297
 Db 1044 GACAACCGCGCGTGTACCGGACCTCTTGGCGGACTTC 1085

RESULT 13
 US-10-425-114-20837
 ; Sequence 20837, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 20837
 ; LENGTH: 1346
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3245-358-A7_FLI
 US-10-425-114-20837

Alignment Scores: 1.03e-09 Length: 1346
 Pred. No.: 154.00 Matches: 78

Percent Similarity: 37.72% Conservative: 48
 Best Local Similarity: 23.35% Mismatches: 120
 Query Match: 9.41% Gaps: 88
 DB: 18 Indels: 16
 10719236-1_232-1164 (1-311) x US-10-425-114-20837 (1-1346)
 Qy 4 ThrValProThrPheSerLeuAlaGluGlnGlnGlyLeuHisGlnAspGluPheArg 23
 Db 228 ACCTCAGCAGCGGTCTCCCTGCGCATGTGACCTTCTCCGCGCGCGAGGTGG 287
 Qy 24 ArgCysLeu-----ArgAspGlyLeuPheTyrLeuThrAspCysGlyLeu 39
 Db 288 CGGCGCATCTCGAGGCGCGGACAGAGATCGCTTCTTCAAGTGTGAACACGCGCTC 347
 Qy 40 ThrAspThrGluLeuSerAlaValAspAlaValIleAspPheGlnHisGlySer 59
 Db 348 TCCCTGAGGCGATGAGGACATGAGAGAGGCTGTGCGAGAGTCTTCCG-----CTG 401
 Qy 60 GluAlaGluValArgAlaValThrSerProValProThrMetArgGlyPheThrGly 79
 Db 402 CCGCGGAGGACAGGCT-----GAG 422
 Qy 80 LeuGluSerGluSerThrAlaGlnIleThr-----AsnThr 91
 Db 423 CTGTACTCGGAGGACACGGGAGGCGGACCGGATCTACTCAGACCATGTTGACAGCG 482
 Qy 92 GlySerTyrSerAspTyrSerMetCysTyrSerMetGlyThrAlaAspAsnLeuPhePro 111
 Db 483 GGGCGGAGAGATGATCGCGGACCTGCTCCGCTGCGCTGCTGCTC-----TTCGCC 533
 Qy 112 Ser---GLYAspPheGluArgIleTrp-----ThrGlnTyrPhe 123
 Db 534 GCCGTGGCGACACGCGCGGCTGCGCGGACCAAGCCCGGAGGTCCGGAGGTGGTG 593
 Qy 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgGluValLeuArg----- 139
 Db 594 GAGAGGTTACAGGTGACAGCGCGGGCTGGGATGAGATCTGCGCTGCTGGAG 653
 Qy 140 AlaThrGlyThrGluProAsp-----GlyGlyValGluAlaPhe 152
 Db 654 GCGCTCGGCTCCGCGCGGACCTGGAGAGGACATCAGCGGGCGG----- 701
 Qy 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGlnHisArgSer 172
 Db 702 -----GACGTGCTCTCCAGCTCAACCACTACCGCGCTGCGCGCGCGCGCG 752
 Qy 173 AlaGluGlnGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192
 Db 753 ACG-----CTGGGCTGCGCGGACCTGTACCGCAACCTCTCACTTGGCTC 800
 Qy 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyGlyAlaPhe 212
 Db 801 -----CTCCCGACAGTGTACTGGC----- 821
 Qy 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223
 Db 822 CTCGAGGTCGCTTACGGGGGCGACTGATCCGGGTGAGCTGTGCGCGCGCTTGGCTC 881
 Qy 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyValValAlaProArgHis 243
 Db 882 GTCAACTTCGGATGCGAGCTTGAGTGTGACGAACGGAGCTTGAAGAGCATGAGAC 941
 Qy 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
 Db 942 CGGGTGATGACCAAC-----CTGGAGTGGCGCGACACACGGTGGCCACG 986
 Qy 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283
 Db 987 TTCATCATGCCACACGAGCTGCTCATCGGC---CCGCGCGCGAGTTCTTCAAGGAC 1043
 Qy 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrp 297

Db 1044 GACAAACCCGCGGTCTACCGCACCTTCACCTTCGGCGACTTC 1085

RESULT 14

US-10-425-114-22268

/ Sequence 22268, Application US/10425114

/ Publication No. US20040034888A1

/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jindong

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Screen, Steven E

/ APPLICANT: Tabaska, Jack E

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

/ FILE REFERENCE: 38-21(5313)B

/ CURRENT APPLICATION NUMBER: US/10/425,114

/ NUMBER OF SEQ ID NOS: 73128

/ SEQ ID NO 22268

/ LENGTH: 1376

/ TYPE: DNA

/ ORGANISM: Zea mays

/ FEATURE:

/ OTHER INFORMATION: Clone ID: LIB357-058-F2_F1

US-10-425-114-22268

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1,076-09	1376	154.00	78	48	120	88	16
Percent Similarity:		37.72%					
Best Local Similarity:		23.35%					
Query Match:		9.41%					

DB: 18 Gaps: 16

10719236-1_232-1164 (1-311) x US-10-425-114-22268 (1-1376)

QY 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGluPheArg 23

DB 179 ACCTCAGGACGAGCTCTCCCTCCATCGAGACCTCTCCCTCGGCGACAGAGTGGCG 238

QY 24 ArgCysLeu-----ArgAspGlyLeuPheThrGlyLeuThrAspCysGlyLeu 39

DB 239 CGCGCATCTCGAGGCGGCGAGAGAGATCGGCTTTCCAGTGGTGAACACGCGCTTC 298

QY 40 ThrAspThrGluLeuLysSerAlaLysAspGlnLeuAlaAspPheGlnHisGlySer 59

DB 299 TCCCTGAGGCGATGCGAGCATGAGACGATGTCGACAGAGTCTTCCG-----CTG 352

QY 60 GluAlaGluLysArgAlaValThrSerProValProThrMetArgGlyPheThrGly 79

DB 353 CGGCGAGACACAGGCC-----GGG 373

QY 80 LeuGluSerGluSerThrAlaGlnIleThr-----AsnThr 91

DB 374 CTGTACTGGAGAGACACGCGGAGGCGGACCGGATCTACTCAGACCATGTTGACACG 433

QY 92 GlySerTyrSerAspTyrSerMetCysTyrSerMetGlyThrAlaAspValLeuPhePro 111

DB 434 GCGCGCGAAGAGTACTGCGCGCATGCTCCCTCGCTGGCTGCTCC-----TTCCCC 484

QY 112 Ser---GlyAspPheGluArgGlyIleThr-----ThrGlnIleThr 123

DB 485 GCGCTCGGCGACGCGCGCGCTGCGCCGACACACCGCGGAGGCTCGGAGGCTGCTG 544

QY 124 AspArgGlnIleThrAlaSerArgAlaValAlaAspGluValLeuArg----- 139

DB 545 GAGAGTTTACGCGTGAACGCGGCGGCTCGGAGATGAGATCTCGCTGCTGCGAG 604

QY 140 AlaThrGlyThrGluProAsp-----GlyGlyValGluAlaPhe 152

DB 605 GCGCTCGGCTCGCGCGCGATACCTGAGAGGGAACATCAGCGCGCG----- 652

QY 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGlnHisArgSer 172

DB 653 -----GACGTGTCCTCCACGTCACACCTACCGCCCTGCGCCGACCCCAAGCC 703

QY 173 AlaGluGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192

DB 704 AGC-----CTGGCCTGCGCGCGCATGTCGACCGGACCTCTCACCCTTGCTC 751

QY 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyAlaPhe 212

DB 752 -----CTCCACGATGATGCTGCG----- 772

QY 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223

DB 773 CTGAGGTGCTCTACAGGCGCATGATCCGGGTGAGCTGTGCGCGCGCTTCTGTC 832

QY 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlnValIleValAlaProArgHis 243

DB 833 GTCACTTCGATGCGACCTTGAAGTTGTGACGAAACGGGATCTTGAAGACATCGAGCAC 892

QY 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerValPhe 263

DB 893 CGGTGATGACCAAC-----CTGGAGTGGCGCGGACCAACGCGTGCACAG 937

QY 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283

DB 938 TTCATCATGCCCCACGAGACTGCTCATCGCG-----CCCGCGCGAGTTCTCAGCGAC 994

QY 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTyr 297

DB 995 GACAAACCCGCGGTCTACCGCACCTTCACCTTCGGCGACTTC 1036

RESULT 15

US-10-425-115-35659

/ Sequence 35659, Application US/10425115

/ Publication No. US20040214272A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa, Thomas J.

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

/ FILE REFERENCE: 38-21(5322)B

/ CURRENT APPLICATION NUMBER: US/10/425,115

/ NUMBER OF SEQ ID NOS: 36326

/ SEQ ID NO 35659

/ LENGTH: 1430

/ TYPE: DNA

/ ORGANISM: Zea mays

/ FEATURE:

/ OTHER INFORMATION: Clone ID: MRT4577_132518C.1

US-10-425-115-35659

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1,136-09	1430	154.00	78	48	120	88	16
Percent Similarity:		37.72%					
Best Local Similarity:		23.35%					
Query Match:		9.41%					

DB: 20 Gaps: 16

10719236-1_232-1164 (1-311) x US-10-425-115-35659 (1-1430)

QY 4 ThrValProThrPheSerLeuAlaGluLeuGlnGlnGlyLeuHisGlnAspGluPheArg 23

DB 208 ACCTCAGGACGAGCTCTCCCTCCATCGAGACCTCTCCCTCGGCGACAGAGTGGCG 267

QY 24 ArgCysLeu-----ArgAspGlyLeuPheThrGlyLeuThrAspCysGlyLeu 39

DB 268 CGCGCATCTCGAGGCGGCGAGAGATCGGCTTTCCAGTGGTGAACACGCGCGCTC 327

```
Qy 40 ThrAspThrGluLeuYrsSerAlaYrsAspIleValIleAspPhePheGluHisGlySer 59
Db 328 TTCCTGAGCGGATGCAAGACATGAGAGCGGTGTGCCAGAGTCTTCGG-----CTG 381
Qy 60 GluValGluYrsArgAlaValThrSerProValProThrMetArgArgGlyPheThrGly 79
Db 382 CCGCGGGAGGACAAAGGCC-----GGG 402
Qy 80 LeuGluSerGluSerThrAlaGlnIleThr-----AsnThr 91
Db 403 CTGTACTCGAGGACACCGGGAGGCGACCGGATCTACTCCAGCACCATGTTCCACACG 462
Qy 92 GlySerTyrSerAspTyrSerMetCysTyrSerMetGlyThrAlaAspAsnLeuPhePro 111
Db 463 GGCGGCGAAGATGACTGCGGACTGCTCGCTCGGCTGCGCTGCTCC-----TTCGCC 513
Qy 112 Ser---GlyAspPheGluArgIleTrp-----ThrGlnTyrPhe 123
Db 514 GCCGTCCGGCGACAGCGCGCGCTGGCCGCAAGCCCGAGAGCTCCGGAGAGTGATG 573
Qy 124 AspArgGlnTyrThrAlaSerArgAlaValAlaArgIleValLeuArg----- 139
Db 574 GAGAGTTTCACGTCGACACCGGGGCTGGGATGGAGATCTCGCGCTGTCGCGAG 633
Qy 140 AlaThrGlyThrGluProAsp-----GlyGlyValGluAlaPhe 152
Db 634 GGCCTCGGCTCCCGCCCTGACTGAGAGGAGCATCAGCGCGGC----- 681
Qy 153 LeuAspCysGluProLeuLeuArgPheArgTyrPheProGlnValProGlnHisArgSer 172
Db 682 -----GACGTGCTCTCCACGTCAACCATACCGCCCTGCGCCGACCCACGCC 732
Qy 173 AlaGluGluGlnProLeuArgMetAlaProHisTyrAspLeuSerMetValThrLeuIle 192
Db 733 ACG-----CTCGGCTGCGCGCATGTGACCGCAACTCTTCACCTTGCTC 780
Qy 193 GlnGlnThrProCysAlaAsnGlyPheValSerLeuGlnAlaGluValGlyAlaPhe 212
Db 781 -----CTCCCCAGCATGTACTGCG----- 801
Qy 213 ThrAspLeuProTyrArg-----ProAspAlaValLeu 223
Db 802 CTCGAGGTGCGCTTACAGGGGCGCATGATCCGGGTGAGCTGTGCCGCGCTTCGTC 861
Qy 224 ValPheCysGlyAlaIleAlaThrLeuValThrGlyGlyGlnValIleValAlaProArgHis 243
Db 862 GTCAACTTCGAGTCCACGCTTGAGGTGTGACGAACGGGATCTTGAGAGCATCGAGCAC 921
Qy 244 HisValAlaAlaProArgArgAspGlnIleAlaGlySerSerArgThrSerSerValPhe 263
Db 922 CCGGTGATGACCAAC-----CTGGAGTGGCGCGGACACACGTTGCCACG 966
Qy 264 PheLeuArgProAsnAlaAspPheThrPheSerValProLeuAlaArgGluCysGlyPhe 283
Db 967 TTCATCATGCCACACACGACTGCTCATTCGAC---CCCGCGCGGAGTTCTCAGCGAC 1023
Qy 284 AspValSerLeuAspGlyGluThrAlaThrPheGlnAspTrp 297
Db 1024 GACAAACCGCGCTGCTACCGGACCTCTACCTTCGCGACTTC 1065
```

Search completed: September 19, 2005, 21:13:51
Job time : 758 secs